Integrated Analysis of Patterning, Morphogenesis, and Cell Divisions in Embryonic Development by in toto Imaging and Quantitative Cell Tracking

Integrated Analysis of Patterning, Morphogenesis, and Cell Divisions in Embryonic Development by *in toto* Imaging and Quantitative Cell Tracking

A dissertation presented

by

Fengzhu Xiong

to

Department of Biological Sciences in Dental Medicine

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological Sciences in Dental Medicine

Harvard University

Cambridge, Massachusetts

September, 2013

Dissertation Advisor: Professor Sean G. Megason Fengzhu Xiong

Integrated Analysis of Patterning, Morphogenesis, and Cell Divisions in Embryonic Development by *in toto* Imaging and Quantitative Cell Tracking

Abstract

Patterning, morphogenesis, and cell divisions are distinct processes during development yet are concurrent and likely highly integrated. However, it has been challenging to investigate them as a whole. Recent advances in imaging and labeling tools make it possible to observe live tissues with high coverage and resolution. In this dissertation work, we developed a novel imaging platform that allowed us to fully capture the early neural tube formation process in live zebrafish embryos at cellular resolution. Importantly, these datasets allow us to reliably track single neural progenitors. These tracks carry information on the history of cell movement, shape change, division, and gene expression all together. By comparing tracks of different progenitor fates, we found they show a spatially noisy response to Sonic hedgehog (Shh) and become specified in a positionally mixed manner, in surprising contrast to the "French Flag" morphogen patterning model. Both cell movement and division contribute to cell mixing. In addition, we decoupled the temporal and genetic regulatory network (GRN) noises in Shh interpretation using tracks that carry both Shh signaling and cell fate reporters. Our tracks suggest that, after specification, progenitors undergo sorting to self-assemble a sharp pattern. Consistent with this hypothesis, we found ectopically induced progenitors move to correct locations. Furthermore, we show that proper adhesion is required for cell sorting to happen (Chapters 2 and 3). In the cleavage stage embryos, the cells on the surface undergo shape changes followed by lineage separation and differentiation. We quantitatively measured this morphogenesis process and tracked cell divisions. By applying a mathematical model we uncover a predictive, and perhaps general link between cell division orientation, mechanical interaction, and the morphogenetic behavior of the whole surface layer (Chapter 4). Finally, we discuss the concepts and tools of cell tracking including a multi-color cell labeling method we developed by modifying the "Brainbow" system (Chapter 5). Together this dissertation showcases the importance and promise of live observation based, quantitative and integrated analysis in our understanding of complex multi-cellular developmental processes.

Contents

Citations to Previously Published Work

Chapter 2 has been previously published in its entirety in this following journal paper:

Xiong, F., Tentner, A.R., Huang, P., Gelas, A., Mosaliganti, K.R., Souhait, L., Rannou, N., Swinburne, I.A., Obholzer, N.D., Cowgill, P.D., Schier, A.F., and Megason, S.G. (2013). Specified neural progenitors sort to form sharp domains after noisy Shh signaling. Cell *153,* 550-561.

Acknowledgements

I would first especially like to thank my advisor Sean for his constant and generous support and guidance during my training. I saw myself grow to be an independent and confident scientist under Sean's mentorship and care. I would like to thank my collaborators Andrea Tentner, Peng Huang and Wenzhe Ma for encouragement, expertise and luck - not to mention great joy - that they brought into my academic life. I also thank Kishore Mosaliganti, Arnaud Gelas, Lydie Souhait, Nicolas Rannou and Hideo Otsuna for their friendship, computational knowledge, assistance, and support. I thank Ramil Noche for teaching me about the amazing model of zebrafish and looking after my study when I joined the lab, Nik Obholzer and Ian Swinburne for friendship, advice and technical expertise that I could always count on, and Dante D'India, Amelia Green, Tom Hiscock, Paul Cowgill, Yao Chen and members in the DePace, Mitchison, Kirschner, Lahav, Schier and Zon labs for their friendship, comments and technical help. I would also like to thank the members of my Dissertation Advisory Committee – Professors Connie Cepko, Tim Mitchison and Alex Schier – for their constant encouragement, insightful comments and wise advice not only in the work described here but also on my academic career in general. I am also especially grateful to Professor Bjorn Olsen and administrators Yulia Pittel, Liz Kervis, and Deb Milstein of the developmental biology department in Harvard School of Dental Medicine, who have encouraged and supported me from my arrival at Harvard to the completion of this dissertation. I also thank my classmates in the graduate program of Biological Sciences in Dental Medicine and the graduate program of Systems Biology for interesting scientific discussions. I am also especially grateful to the administrative staff in the department of Systems Biology for help and support. Second to last but not least, I would like to thank my friends and colleagues in Harvard and the research communities of zebrafish, imaging and developmental biology worldwide, who gave me tools, help, ideas and inspirations every now and then during my scientific queries. Finally, I would like to thank my family and friends for their unwavering support that helped me complete this dissertation.

Dedicated to my loving wife Mengqi, and my parents Mingmei and Hongbin.

Chapter 1

Integrated Analysis of Patterning, Morphogenesis and Cell Divisions in Embryonic Development by *in toto* Imaging and Quantitative Cell Tracking - An Introduction

Fengzhu Xiong, Sean G. Megason

Department of Systems Biology, Harvard Medical School, Boston, 02115, USA

Author Contributions: F.X. reviewed the literature and wrote the manuscript under the guidance of S.G.M.

Development relies on cell-cell interactions to produce spatial-temporal organizations of gene expression (patterning) and form (morphogenesis). These interactions are carried out by molecules that cells produce and respond to, such as diffusible signals, membrane bound receptors and adhesion junction proteins. Together they generate a limited range of cell behaviors including target gene expression, movement, growth, shape change, apoptosis and division. The collective behavior of the cell population then leads to developmental phenotypes of great variety, from stripes of a zebra to orthogonal semicircular canals in the inner ear.

A lot of the aforementioned molecules have been discovered and characterized through genetics and biochemistry, and altering them *in vivo* has been powerful in tearing apart the black box of development. However, much less is understood about how cell behaviors mediated by these molecules produce the observed phenotypes, either in normal or perturbed conditions for several reasons. First, the mechanism by which a molecule changes cell behavior within a cell is often

not clear and highly context dependent. Second, phenotypes arise on much longer space (across many cells) and time scales than biochemical reactions making it difficult to determine when the molecule is influential and which cells matter most. Third, heterogeneity and noise are widespread in cell populations. Fourth, distinct types of cell behaviors could lead to a similar result. For example, a boundary of gene expression domains could arise from a threshold of a signaling molecule, community effect, cell sorting, chemotaxis or death of ectopic cells. Fifth, it is not only the "final" phenotype, but also the speed, robustness, cost and flexibility of the developmental processes leading to the phenotype that are under evolutionary pressure. How might different cell behaviors act together to promote these properties of the system?

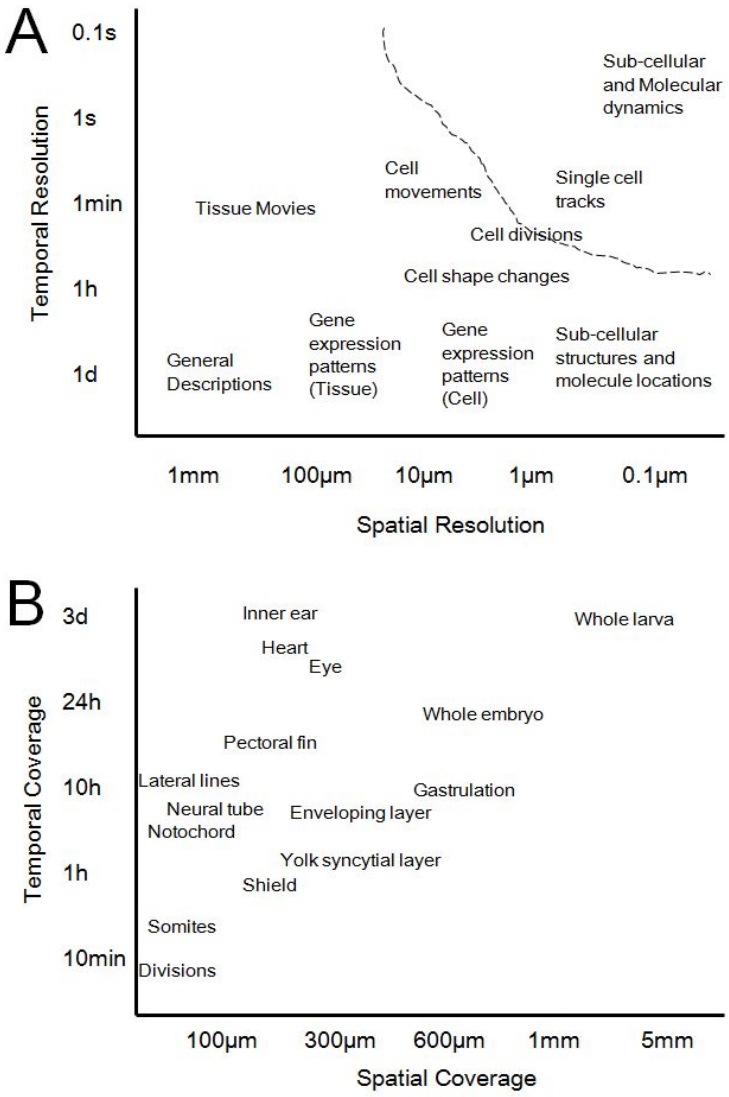


Figure 1.1. Information vs. spatial-temporal resolution and coverage in imaging

(A) Resolution of imaging determines the type of information that can be learned about developmental processes, this plot assumes zebrafish as the model system. Cell tracking is possible to the top right of the dashed line.

(B) Coverage of imaging required by different developmental processes, tissues or organs, assuming zebrafish as the model system.

Recent advances in live imaging techniques offer new opportunities to answer these questions

(Megason and Fraser, 2007). At a spatial-temporal resolution that is sufficient for cell tracking, imaging becomes very powerful (Figure 1.1A). Cell tracks carry information of movement, shape change, division and gene expression on the same time axis, making it possible to analyze the exact sequence of these cellular events, as well as how groups of cells may carry them out differently or cooperatively leading up to the pattern or form. For cell tracking to be meaningful on the tissue scale, it is important to have a corresponding spatial-temporal coverage to allow a number of cells (or ideally, all cells in the tissue of study) to be tracked until the events of interest have occurred (Figure 1.1B). We have coined the term "*in toto*" imaging to describe this type of observation (Megason and Fraser, 2003). For example, Keller et al. captured the cell dynamics in whole early zebrafish embryos using a lightsheet microscope (Keller et al., 2008). By comparing a wild type and an MZ*oep* mutant which lacks mesoderm (Gritsman et al., 1999), the authors identified a key "embolic wave" stage during which many cells internalize, a process that is greatly reduced in the mutant. The full coverage imaging and cell tracks allowed the authors to determine exact number of affected cells and the magnitude of mutation effect on cell and tissue movements. Another example comes from Martin et al., who used high frequency imaging to analyze actomyosin contractions that drive ventral furrow formation (Martin et al., 2009). The authors show that the contractions come as pulses differentially regulated by Snail and Twist. The ability to increase spatial temporal resolution to allow measurement of cell shape and molecular dynamics in individual cells was essential for the authors to arrive at an elegant "ratchet" model for this process. These examples show that, by merely imaging better, it is possible and often the case to not only improve the quantitativeness and accuracy of current models, but also uncover novel mechanisms that are additional or contrary to current models.

On the other hand, these imaging data may help theorists initiate an attack on the so-called "last

refuge of the mathematically incompetent biologists" (Gilbert 2000) - another name for developmental biology that highlights its fruitful prospect if quantitative formulations are made possible in this field. Mathematical modeling logically constructs hypotheses that describe how things work and can simulate complex developmental systems *in silico*. While practical modeling still relies on data for formulation and validation, it can go beyond the data to overcome several limitations of wet experiments. Firstly, using modeling, one may be able to infer and test mechanisms more specifically at a reasonable degree of perturbation (Barkoulas et al., 2013) that allows analysis of physiological responses, instead of introducing perturbations that overwhelm the system; Secondly, it allows a comparison between different interactions and the isolation of important ones - among all that have been discovered. For example, 'Core' motifs of gene regulatory networks are found, and validated, by probing the possible parameter space that define the interactions between players (Ma et al., 2009); Finally, when experimental data is of good quality, modeling may be able to predict an exact mechanism, or existence of an unknown player thereby leading to directed discoveries (Hodgkin and Huxley, 1952; Neher and Sakmann, 1976). Such cases are rarer but remain an exciting possibility. With the advance of imaging discussed above, one may speculate audaciously that modeling will gain increasingly stronger predictive powers in developmental biology.

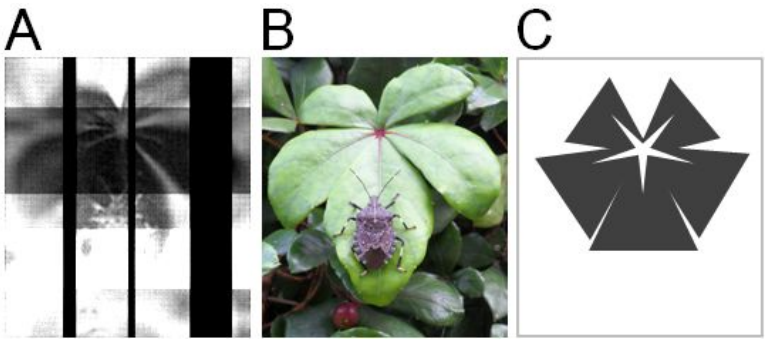


Figure 1.2. Imaging, modeling and the real picture

Figure 1.2 (*Continued*) (A) Bottom-up method of imaging, capturing as much detailed information as possible but still far from the real picture.

(B) The real picture. Photo credit: F. Xiong.

(C) Top-down method of modeling, using simplifying assumptions and principles to understand the key features of the picture.

Together, imaging and modeling bring us closer to the real picture from different (and arguably opposite) directions (Figure 1.2). Modeling helps imaging stay focused on important principles and not going into trivial details just for the sake of resolution and coverage, while imaging helps modeling establish solid grounds and stick to reality instead of wandering off into science fiction or over-simplifications that are not useful. In the following chapters of this dissertation, we will demonstrate examples of high resolution live imaging in the zebrafish neural tube (chapter 2,3) and early embryonic surface (chapter 4). We optimized many technical variables to achieve feasible balances between resolution and coverage that allowed us to fully capture early neural tube formation and EVL (enveloping layer) morphogenesis on the embryonic surface at cellular resolution. Using cell tracking we established gene expression dynamics, lineage trees, cell shape evolution and cell movement to assess the interactions between tissue patterning and morphogenesis in a systematic and quantitative manner. We also use imaging data to formulate and test a model of cell shape change under the influence of division patterns (chapter 4). In addition, we show a design of multi-color cell tracking method in zebrafish (chapter 5) based on modifications from the brainbow (Livet et al., 2007) system. These projects span four and a half years proportioned approximately as: chapter 2,3 - 60%, chapter 4 - 30%, chapter 5 - 10%.

ACKNOWLEDGMENTS

The authors thank C. Cepko, A. Schier, T. Mitchison and Megason lab members for helpful discussions. This work was supported by R01 DC010791 from NIH. F.X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

CHAPTER 1 REFERENCES

Barkoulas, M., van Zon, J.S., Milloz, J., van Oudenaarden, A., and Felix, M.A. (2013). Robustness and epistasis in the C. elegans vulval signaling network revealed by pathway dosage modulation. Dev. Cell. *24*, 64-75.

Gilbert SF. Developmental Biology. 6th edition. Sunderland (MA): Sinauer Associates; (2000). Mathematical Modeling of Development. Available from: http://www.ncbi.nlm.nih.gov/books/NBK10126/

Gritsman, K., Zhang, J., Cheng, S., Heckscher, E., Talbot, W.S., and Schier, A.F. (1999). The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. Cell *97*, 121-132.

Hodgkin, A.L., and Huxley, A.F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiol. *117*, 500-544.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322*, 1065-1069.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56-62.

Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. Cell *138*, 760-773.

Martin, A.C., Kaschube, M., and Wieschaus, E.F. (2009). Pulsed contractions of an actin-myosin network drive apical constriction. Nature *457,* 495-499.

Megason, S.G., and Fraser, S.E. (2007). Imaging in systems biology. Cell *130,* 784-795.

Megason, S.G., and Fraser, S.E. (2003). Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. Mech. Dev. *120,* 1407-1420.

Neher, E., and Sakmann, B. (1976). Single-channel currents recorded from membrane of denervated frog muscle fibres. Nature *260,* 799-802.

Chapter 2

Specified Neural Progenitors Sort to Form Sharp Domains after Noisy Shh Signaling

Fengzhu Xiong[1], Andrea R. Tentner[1], Peng Huang[2], Arnaud Gelas[1], Kishore R. Mosaliganti[1], Lydie Souhait[1], Nicolas Rannou[1], Ian A. Swinburne[1], Nikolaus D. Obholzer[1], Paul D. Cowgill[1], Alexander F. Schier[2], Sean G. Megason[1]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA

[2]Department of Molecular and Cellular Biology, Center for Brain Science, Harvard Stem Cell Institute, Broad Institute, Harvard University, Cambridge, MA 02138, USA

Author Contributions: F.X. and S.G.M. conceived this study. F.X., A.R.T. and P.H. performed the experiments. F.X. analyzed the data. A.G., K.R.M., L.S. and N.R. provided tools and assistance for data analysis. P.H., I.A.S., N.D.O., P.D.C, and A.F.S. provided reagents and technical assistance. F.X. and S.G.M. wrote the manuscript. All authors discussed the manuscript and contributed to writing.

SUMMARY

Sharply delineated domains of cell types arise in developing tissues under the instruction of inductive signal (morphogen) gradients, which specify distinct cell fates at different signal levels. The translation of a morphogen gradient into discrete spatial domains relies on precise signal responses at stable cell positions. However, cells in developing tissues undergoing morphogenesis and proliferation often experience complex movements, which may affect their

morphogen exposure, specification, and positioning. How is a clear pattern achieved with cells moving around? Using in toto imaging of the zebrafish neural tube, we analyzed specification patterns and movement trajectories of neural progenitors. We found that specified progenitors of different fates are spatially mixed following heterogeneous Sonic Hedgehog signaling responses. Cell sorting then rearranges them into sharply bordered domains. Ectopically induced motorneuron progenitors also robustly sort to correct locations. Our results reveal that cell sorting acts to correct imprecision of spatial patterning by noisy inductive signals.

HIGHLIGHTS

- Novel in toto imaging platform for watching neural tube formation
- Neural progenitors are specified in a “salt and pepper” pattern
- Sharp domains and boundaries of gene expression form by cell sorting
- Ectopic progenitors robustly form sharp domains similar to the normal pattern

INTRODUCTION

Two central questions in developmental biology are how cell type diversity is generated, and how these types are organized into patterns of structural and functional significance. The classic “French Flag” model (Wolpert, 1969) put forward the idea of morphogen patterning that mechanistically couples specification and spatial arrangement. In this view, a gradient of a diffusible signal across a field of naive cells defines spatial domains of cell types between concentration thresholds. Recent studies have challenged and extended this model in several aspects: First, a signaling gradient may not be sufficient to generate precise cell type boundaries, given the noise inherent in molecular processes and the limited information content of gradients

in vivo (Paulsson, 2004; Lander et al, 2009). Second, the timing of exposure to the signal, in addition to concentration, contributes to cell fate choices (Ahn and Joyner, 2004; Harfe et al, 2004; Dessaud et al, 2007). Third, the position of a cell relative to a morphogen source may change in time through cell migration and division (Kay and Thompson, 2009). Finally, lateral cell-cell interactions such as cell sorting may also be involved in boundary formation (Lawrence et al, 1999; Nicol et al, 1999; Xu et al, 1999).

A prominent example of morphogen patterning is the vertebrate ventral neural tube. In this system, sharply bordered progenitor domains form along the ventral-dorsal axis (Jessell, 2000; Figure 2.2A). This spatial arrangement is important for the localization, migration, and wiring of neurons born from these domains (Lewis and Eisen, 2003; Sürmeli et al, 2011). Significant molecular insights have been generated towards the understanding of how this pattern forms: First, the secreted signaling protein Sonic Hedgehog (Shh) is produced in the notochord underlying the neural tube and later in the floor plate (Krauss et al, 1993; Echelard et al, 1993), and likely forms a ventral to dorsal gradient in the neural tube (Yamada et al, 1993; Chamberlain et al, 2008). Second, gene expression induced by different Shh signaling levels as a function of concentration and duration of exposure *in vitro* parallels the spatial ordering of the expression domains of the same genes *in vivo* (Ericson et al, 1997; Dessaud et al, 2007). Third, intracellular gene regulatory network (GRN) interactions between Shh regulated transcription factors establish stable and discrete fates that no longer depend on Shh (Lek et al, 2010; Balaskas et al, 2012). Together, these studies provide the molecular scenario of morphogen patterning in the neural tube: each cell measures its Shh exposure and enters a corresponding state of gene expression; the states dynamically evolve under the GRN to become self-sustaining, mutually exclusive, and cell type specific; the Shh gradient is thus translated into discrete progenitor

domains. In this model, the shape of morphogen gradient in time and space is directly predictive of the final pattern. Therefore, for the sharply bordered spatial domains in the neural tube to form, Shh exposure levels as a function of position and time must be precise, especially at the putative domain boundaries. In addition, cells should maintain stable positions relative to the source of Shh to receive a correct signal input over time. It is unclear if these requirements for low signaling and positional noise are found *in vivo*, or if additional mechanisms are required to ensure the precision of patterning.

The dynamics of cell movements might provide an answer to these questions. The transition of the neural plate to the neural tube involves extensive cell migration, intercalation, and proliferation (Schoenwolf, 1991; Kimmel et al, 1994; Ciruna et al, 2006) that take place concurrently with Shh gradient formation and interpretation (Marti et al, 1995; Takamiya and Campos-Ortega, 2006). Studies using clonal labeling show cell mixing during morphogenesis and after divisions to variable degrees in the neural tubes of different vertebrates (Leber et al, 1995; Erskine et al, 1998; Inoue et al, 2000; Park et al, 2004), depending on the developmental stage and the anterior-posterior (AP) level. These cellular positional dynamics may affect patterning in several ways. First, movement of Shh producing and responding cells may alter the spatial distribution of Shh ligands among the progenitors, affecting the morphogen gradient. Second, movement of an unspecified progenitor in the gradient may cause its Shh exposure level to change over time, potentially affecting its fate decision (Dessaud et al, 2007). Finally, movement of specified progenitors may either disrupt or sharpen domain boundaries. To evaluate these possibilities, it is essential to understand how individual progenitors behave throughout patterning, proliferation, and morphogenesis.

Here, we use *in toto* imaging to fully capture ventral neural tube formation with single cell

resolution in living zebrafish embryos and report systematic cell tracking analysis of the movies. Our results reveal that intensive cell movements accompany patterning. Shh responding cells show spatial heterogeneity of signaling and become specified to different ventral fates in intermingled distributions. Surprisingly, they then sort out into sharply bordered domains in a robust and Shh independent manner to make the final pattern. Cadherin-mediated cell adhesion is required for the sorting process. These data support a revised "French flag" model where pattern formation in the neural tube is achieved by sorting of specified cells following noisy morphogen-based specification.

RESULTS

In toto imaging reveals cell dynamics during neural tube patterning in zebrafish

The lack of understanding of neural progenitor movements is mainly due to the unavailability of live cell tracking data. Direct imaging of the neural plate is challenging as it undergoes drastic morphogenetic movements, including transition of a horizontal lateral-medial (LM) axis to the vertical dorsal-ventral (DV) axis, morphological and polarity changes of cells, and frequent cell divisions (Clarke, 2009). We designed a novel *in toto* imaging (Megason and Fraser, 2003) system in zebrafish embryos, whose fast development, small size and transparency make full-coverage live imaging feasible. Using an immersed dorsal mount, we allow unrestricted morphogenesis while the embryo sits stably in the field of view (Figure 2.1A, 2.2B). This enables uninterrupted imaging sessions on single embryos from early neural plate to neural tube stages using confocal/2-photon microscopy. We acquired high resolution image stacks every 2-3 minutes of healthy embryos labeled with nuclear/membrane fluorescent proteins and transgenic reporters (Figure 2.1B, 2.2B,C,E, data not shown). These data provide the first trackable movies

of ventral neural tube formation (Movie S2.1). They thoroughly cover the period of Shh expression, progenitor responses, and the establishment of stable pattern (Figure 2.1D.i, Krauss et al, 1993; Huang et al, 2012), allowing us to directly watch patterning (Figure 2.1E, Movie S2.2). We manually tracked cells using the GoFigure2 software that we have developed (Figure 2.1C, 2.2D, Extended Experimental Procedures). These tracks provide systematic and quantitative data on transgenic reporter expression (Figure 2.1D.ii), lineage relationships (Figure 2.1D.iii), and importantly, positional dynamics (Figure 2.1D.iv) of the neural progenitors, allowing us to study the role of cell movements in pattern formation.

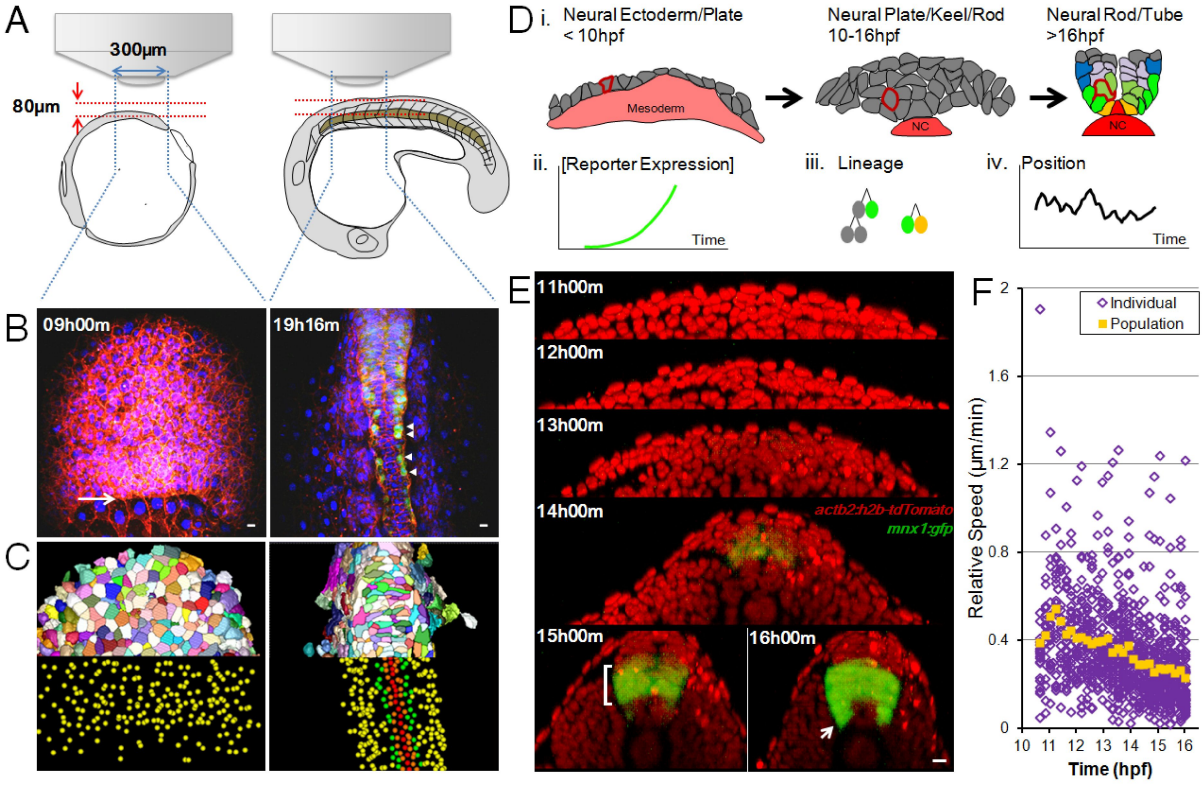


Figure 2.1. *In toto* imaging captures dynamics of neural progenitors during neural tube formation in zebrafish embryos

(A) Schematic illustration of imaging set-up. See also Extended Experimental Procedures.

(B) Sample time points of raw data rendered in 3D projection dorsal view. Red: mem-citrine,

Figure 2.1 (*Continued*) Blue: h2b-cherry, Green: *mnx1:gfp*. Arrow: the frontier of epiboly movement. Arrowheads: differentiating motorneurons (MNs). All time annotations are hours (and minutes) post fertilization (hpf). All scale bars: 10μm. See also Movie S2.1.

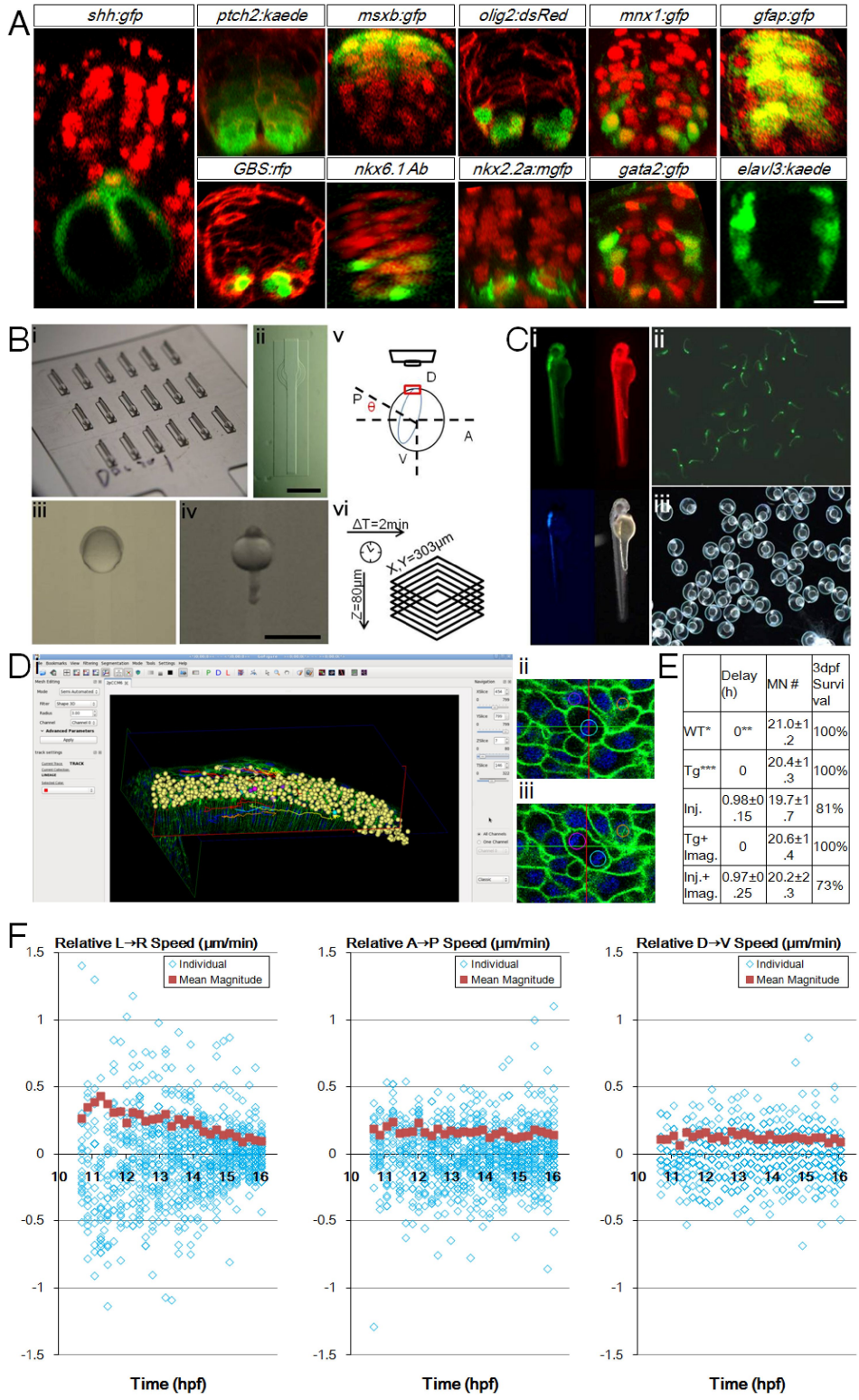
(C) Processed data by GoFigure2 and ACME (Mosaliganti et al, 2012) software from images in B. Top halves: membrane segmentations (random colors to distinguish neighbors); Bottom halves: nuclei segmentations for cell tracking (Red: medial floor plate (MFPs). Orange: lateral floor plate (LFPs). Green: motorneuron progenitors (pMNs). Yellow: unidentified cells).

(D) Schematic illustration of cell tracking analysis. Drawings are based on cross-section images, colors are assigned based on marker expression (Red: *Shh*, Yellow: *nkx2.2a*, Green: *mnx1*, Blue: *gata2*). Part of the notochord (NC, *Shh*+) is included. (i). Morphogenesis during the patterning process, single cells can be tracked throughout (e.g. highlighted cell with red membrane). Tracks carry information of reporter expression (ii), lineage relationships (iii) and movement trajectories (iv). See also Figure 2.2A.

(E) Cross-sectional view (Dorsal side up) of sample dataset. Red: nuclei. A GFP+ “stripe” domain emerges (bracket, bottom left image). Arrow: Differentiating MNs exiting the GFP domain. See also Movie S2.2.

(F) Relative speed of cell movement during neural tube formation. Each purple mark represents the speed of a single cell, 41 tracked ventral cells are plotted. Relative speed is calculated by dividing a cell’s positional change (μm) between 2 time points over the time difference (11.5 minutes). Position is measured relative to the average position of all tracked cells to eliminate global movements introduced by embryo rotation/shifting. Orange marks: average speed. See also Figure 2.2F.

To assess the extent of cell movement, we calculated progenitor speeds at different times (Figure 2.1F, 2.2F). Cells show extensive movements which slow down gradually on the population level as the neural tube forms between 10 and 16 hours-post-fertilization (hpf). For individual cells, mobility is reduced when they become epithelialized (Movie S2.3, data not shown). These data indicate that patterning occurs at a time when cells are moving, on average at a fast rate of one cell diameter every 10-20 minutes, not when the field of cells is static. It is intriguing that sharp spatial domains arise correctly in such a dynamic environment.



	Delay (h)	MN #	3dpf Survival
WT*	0**	21.0±1.2	100%
Tg***	0	20.4±1.3	100%
Inj.	0.98±0.15	19.7±1.7	81%
Tg+ Imag.	0	20.6±1.4	100%
Inj.+ Imag.	0.97±0.25	20.2±2.3	73%

Figure 2.2. Zebrafish ventral neural tube pattern, in toto imaging set-up, controls and analysis tools (related to Figure 2.1)

Figure 2.2 (*Continued*) (A) Gene expression patterns of 24hpf ventral neural tube. Red signals are ubiquitous membrane or nuclear markers to highlight cell locations and tissue morphology. Consistent with previous studies (Park et al, 2002; Schafer et al, 2005; Kimura et al, 2008; Huang et al, 2012), a single column of triangular (apically constricted) medial floor plate (MFP, red color in Figure 2.1Di) cells (*shh:gfp*+) sits in the ventral center bordering the dorsal top center of the notochord. 2 columns of lateral floor plate (LFP, yellow in Figure 2.1Di) cells (*nkx2.2a:mgfp*+, *olig2:dsRed*+(weak)) flank the MFP domain on both sides. In the LFP domain some cells become Kolmer-Agduhr (KA") neurons (*gata2:gfp*+). Dorsal to the floor plate cells sit the motorneuron progenitors (pMNs, light green in Figure 2.1Di) and motorneurons (MNs, green in Figure 2.1Di) cells (*olig2:dsRed*+, *mnx1:gfp*+(strong)). Further dorsal to the 2-3 cell thick pMN domain sit the p2 and V2 interneuron (blue in Figure 2.1Di) cells (*gata2:gfp*+, *mnx1:gfp*+(weak)). Weak *mnx1:gfp* expression extends beyond *nkx6.1*+ domain to 2/3 of the whole tube, and is bordered dorsally by *msxb:gfp*+ cells. At this stage, a *ptch2:kaede* gradient is seen across the DV axis, while GLI reporter *GBS:rfp*+ cells are mostly ventral. Differentiated neurons (*elval3:kaede*+) sit on the lateral sides away from the apical glia and progenitor cells (*gfap:gfp*+). Injections: membrane-citrine/cherry/RFP for *ptch2*, *olig2*, and *GBS*(GLI binding site) embryos; h2b-cherry for *shh*, *msxb*, *mnx1*, *gfap*, *nkx2.2a*, *gata2* embryos. DAPI staining: *nkx6.1* embryos. Scale bar: 10μm.

(B) Imaging set-up used in the study. (i) Lucite dorsal mount template and (ii) cast agarose mount (for a full description, see Megason 2009). (iii) Embryo mounted at 6hpf, dorsal side (shield) is facing up in the middle. (iv) Embryo mounted at 24hpf, dorsal side (neural tube) is facing up in the middle. (v) Mounting angle for high resolution timelapse imaging of the neural plate. 7-9hpf embryos are used with the vegetal axis at an angel θ with the horizontal plane, θ

Figure 2.2 (*Continued*) depends on epiboly progress of the embryo. Note that with this set-up uncertainties still remain that the embryos may rotate or tilt that cause the effective cell tracking time window to be cut short. In this report 26% of movies captured (13/50) provide optimal coverage and signal for thorough cell tracking and 80% (40/50) contain useful lineage and morphological information (for a full description, see Extended Experimental Procedures). (vi) Typically used timelapse parameters. Scale bars: 1mm.

(C) Health control of recovered embryo from timelapse imaging. (i) Day 2 embryo shows normal morphology. Data from embryos that showed abnormality at day 2 were not included in analysis. Red: h2b-cherry. Green: mem-citrine. Blue: *mnx1:gfp*. (ii, iii) Progeny of raised embryo from (i). Green signal: *mnx1:gfp*. Most recovered embryos were raised to 5 days and euthanized after confirming normal morphology and swimming behavior.

(D) Data analysis by GoFigure2. (i) User interface showing segmentation and tracking window of neural tube cells. (ii, iii) Example division tracked with spherical segmentation (for a full description, see Extended Experimental Procedures).

(E) Experimental effects on developmental parameters. Injection of high concentration of labeling mRNAs can cause apparent delay for epiboly onset and progress, but does not alter the number of MNs at 28hpf or the neural tube progenitor domain pattern (data not shown). Imaging does not affect developmental parameters. *WT refers to *tg(mnx1:gfp)* embryos that do not have *actb2* transgenes and are uninjected, n=9. **Timing estimated by epiboly progress, setting WT group as 0. ***Tg refers to *tg(actb2:mem-citrine/actb2:h2b-cherry)* in addition to *tg(mnx1:gfp)*, uninjected, n=6. Injections are 2.3nl 30-40ng/µl each mem-citrine and h2b-cherry mRNA at single cell stage, into the cell or the yolk center, n=16. Tg+imaging, n=5. Injection+imaging, n=11.

Figure 2.2 (*Continued*) (F) Cell movement speeds on different embryonic axes. Speeds are decomposed into vectors from Figure 2.1F, negative values indicate moving in opposite direction as indicated by the arrow in the panel title. Population average values are calculated from the absolute magnitude values of individual speeds regardless of direction. These plots indicate lateral-medial movement is the main component of speed change in Figure 2.1F. The DV speeds show more overlap of data points due to lower resolution in the axial dimension (closely aligned with DV axis of the embryo) of the raw datasets.

Shh expressing and responding cells show dynamic movements and heterogeneous levels

To characterize how patterns of Shh signaling may change during the cell movements, we first imaged Shh reporter *tg(shh:gfp)* (Shkumatava et al, 2004) embryos to follow Shh producing cells (Figure 2.3A). Shh expression begins early during epiboly, before there is a notochord or neural tube (Krauss et al, 1993). At this stage, pre-notochord axial mesoderm cells form a wide *shh:gfp+* plate underneath the neural ectoderm, rendering >10 future neural plate cells in cross section as direct neighbors to Shh producing cells. This arrangement changes drastically as the notochord condenses and medial floor plate (MFP) cells start to express Shh, until finally only 2 neural tube cells directly border Shh producing cells (MFPs).

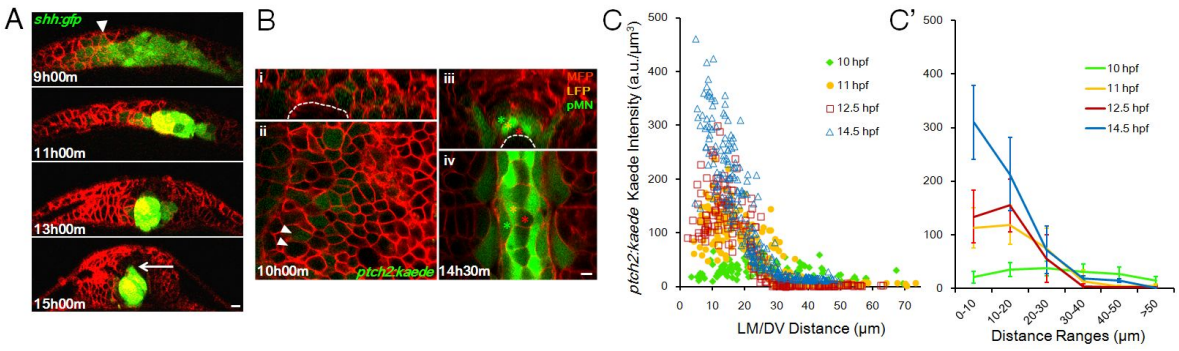


Figure 2.3. Shape changes of Shh gradient and heterogeneity in spatial distribution of

Figure 2.3 (*Continued*) **responses**

(A) Time course of notochord formation by *shh:gfp*+ cells in cross-section. Red: mem-mCherry (same below). Arrowhead: GFP+ cells in neural ectoderm/plate. Arrow: MFP cell expressing GFP. All scale bars: 10μm.

(B) Cross-section (i, iii) and longitudinal-section (ii, iv) of *ptch2:kaede* expression pattern. Arrowheads: Neighbor cells with different Kaede levels. Asterisks: Stereotypic cell fates at the indicated locations. See also Figure 2.2A.

(C,C') Kaede level spatial distribution through time. Each mark represents a segmented cell with measured position and fluorescence intensity. C': spatially averaged (±s.d.) representation of C. Kaede intensities in the notochord cells were subtracted as background.

These dynamic movements of source cells may generate a spatially and temporally changing Shh signaling profile. For example, cells might receive less Shh after moving away from Shh producing cells that are initially their neighbors. To explore this idea, we imaged *tgBAC(ptch2:kaede)* (Huang et al, 2012) embryos which report the level of Shh signaling in responding cells (Figure 2.3B, 2.3A). Interestingly, at 10hpf, neighbor cells at the same location often have very different Kaede levels, and some Kaede+ cells can be found at large distances from the notochord resulting in a highly heterogeneous spatial response distribution (Figure 2.2B.i,ii). The heterogeneity persists as the neural keel forms (12hpf, Movie S2.4), and lasts until after 14.5hpf at which time a clear and sharp gradient can be seen corresponding to different stereotypic cell type locations (Figure 2.3B.iii,iv, Movie S2.4). To compare Shh spatial responses across different neural plate/tube morphologies, we measured cell positions and reporter intensities by GoFigure2 (Figure 2.4A-E). The quantification (Figure 2.2C,C') confirms direct

observations from images, and further indicates that at 10hpf, the Shh response gradient is heterogeneous, broad, and almost flat over a 60μm range. It gradually becomes steeper and less heterogeneous over time as Kaede levels increase in the 30μm range and drop beyond. These data show that each position has a different temporal Shh response profile that is further varied due to local heterogeneity, likely modulated by the movement of both source and responding cells. Together, our observations pose a challenge to the positional specification model where conceptually, a static progenitor field and a smooth morphogen gradient are required for precise pattern formation. How do neural progenitors get patterned correctly into "stripes" (Figure 2.1E, 2.2A) when neither a static field nor a smooth gradient exists?

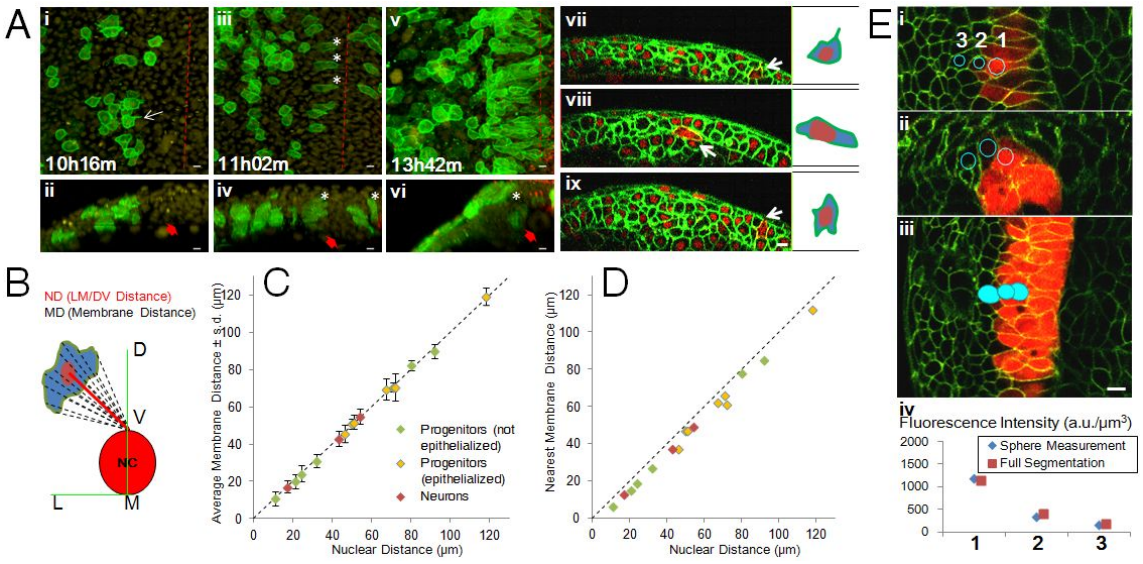


Figure 2.4. Morphological changes of moving progenitors, measuring distances to patterning land marks and marker intensities (related to Figure 2.3)

(A) Mosaic membrane labeling analysis for cell morphology. (i to vi) *Tg(actb2:h2b-tdTomato)* embryo with one cell injected by *mem-citrine* mRNA at 16-cell stage. (i, iii, v) top views, (ii, iv, vi) corresponding cross-section views. White arrow in (i): projection of the cell. Asterisks: epithelialized cells. Filled red arrows: location of the notochord. Dashed lines: midline. (vii to ix)

Figure 2.4 (*Continued*) Embryos injected with *mem-citrine* and *h2b-cherry* mRNAs at one cell stage and further injected with *mem-mCherry* mRNA at 32 cell stage. White arrows: mem-mCherry+ cells. Cartoons on the right highlight morphology. At this resolution the projections of single cells rarely extend beyond cell diameters (i, vii, data not shown, n>100). The cells become epithelialized as they enter the neural keel, the direction of stretching is roughly perpendicular to the LM/DV (Lateral-medial/Dorsal-ventral) axis (ii, iv, data not shown). All scale bars: 10μm. See also Movie S2.3.

(B) LM/DV distance compared to membrane distance. The top point of the notochord on the midline is used as the landmark for all positions in this study, defined as LM/DV distance. As both neural plate/tube and the notochord undergo dramatic morphological changes in the time windows studied, the LM/DV axis provides a suitable reference system. Note that LM/DV distance does not represent or imply Shh signal concentration (Further addressed in Figure 2.3).

(C) Distance measurement by nucleus compared to integrated measurement from membrane points. Nuclear distance is measured as the distance from the nuclear segmentation center to the reference point (panel B). To measure the integrated membrane distance, 20 points were selected along the membrane following evenly changing angels and their distances to the reference are averaged and plotted ±s.d.

(D) Similar plot as (C) plotting minimum membrane distance (closest point of the cell membrane to the reference). These plots show that nuclear distance to reference (LM/DV distance) is an accurate approximation of cell position regardless of cell type and morphology.

(E) Intensity measurement of reporter expression in GoFigure2. Spherical meshes are generated by manually clicking to place a seed inside a cell. The membrane signal provides confinement for verification in different views (top images) if the sphere is completely inside the cell. The

Figure 2.4 (*Continued*) sum of intensity of every voxel inside the sphere is divided by the volume of the sphere to generate the intensity reading used in this study. A more laborious method to measure intensity is to generate full segmentation (bottom image) using the membrane signal. The results from both methods differ by less than 5% (bottom plot, data not shown) and sphere measurement is sufficient to distinguish cells with slight differences (e.g. Cell#2 and #3).

Progenitors make early fate decisions in wide and overlapping ranges

To characterize the spatial distribution and timing of specification of the progenitors, we tracked the motorneuron progenitors (pMNs) and the lateral floor plate cells (LFPs), using *mnx1:gfp*, *olig2:gfp* and *nkx2.2a:mgfp* expression to distinguish their fates (Jessen et al, 1998; Shin et al, 2003; Flanagan-Steet et al, 2005; Figure 2.2A). Previous studies have detected pMN and LFP marker expressions before 12hpf (Korzh et al, 1993; Schafer et al, 2005), suggesting early specification in the neural plate stages. To capture the earliest *mnx1:gfp*+ cells, we performed imaging without other fluorescent cell markers (Figure 2.5A). Scattered GFP+ cells can be found shortly after 10.5hpf, followed by more in a wide LM range, and interestingly, intermingled with GFP- cells. In trackable datasets where ubiquitous cell markers are used, *mnx1:gfp*+ cells can only be distinguished later (13hpf) because of bleed-through signal, but importantly, these tracks show that GFP levels increase monotonically and LFP cells do not turn on GFP (Figure 2.5B). Moreover, this GFP increase is unaffected by Shh inhibitor Cyclopamine (Figure 2.6A), suggesting independence of GFP expression from further Shh signaling. These data indicate that the onset of *mnx1:gfp* marks actual fate specification to pMN instead of LFP. To further assess timing of pMN specification, we performed a time course treatment of Cyclopamine and counted MN numbers from treated embryos (Figure 2.5C, 2.6B). Early blockade (before 12hpf) of Shh

activity greatly reduces MNs, whereas later treatment causes a milder reduction, suggesting a significant portion of pMNs are specified early.

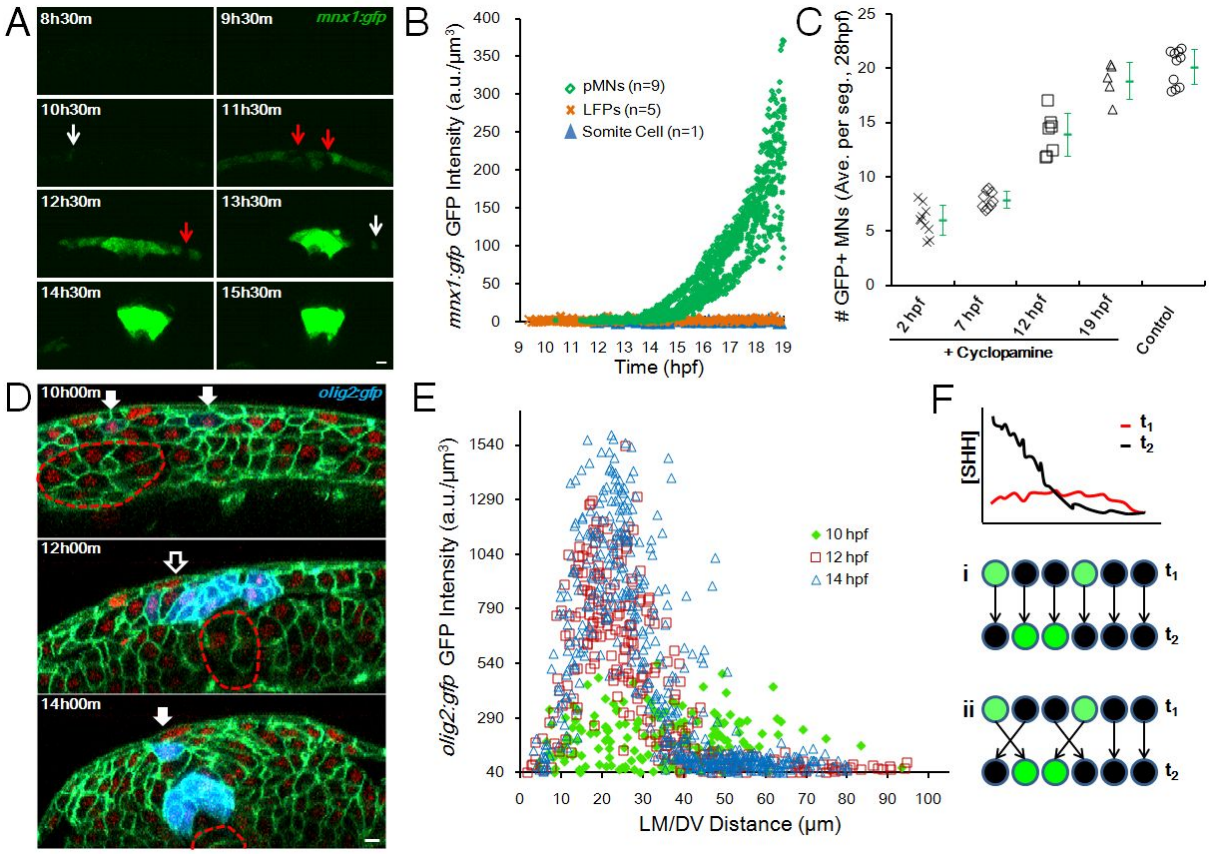


Figure 2.5. Progenitor fates are specified during cell movements in mixed distributions

(A) Time course of *mnx1:gfp* expression. Images are cross-sectional examples. Red arrows: mixed negative cells. White arrows: scattered positive cells. All scale bars: 10μm.

(B) GFP (*mnx1:gfp*) levels in tracked cells through time. See also Figure 2.6A.

(C) Time course of Cyclopamine inhibition of pMN specification. Treatment of 100μM Cyclopamine started at indicated times and MNs were counted at 28hpf as an indicator of pMN number. Numbers are averaged per embryo by number of neural segments counted. Green marks: average (±s.d). See also Figure 2.6B.

(D) GFP(*olig2:gfp*) domain formation. Green: cell membrane. Red: cell nucleus. Filled arrows:

Figure 2.5 (*Continued*) scattered positive cells. Empty arrows: mixed negative cells. Dashed lines: notochord boundary.

(E) Spatial distribution of *olig2:gfp*+ cells. At 10hpf they scatter in a wider range and are mixed with negative cells, in contrast, at 14hpf positive cells form a major “stripe” between 15 and 30μm where negative cells are absent.

(F) Two models for sharp stripe formation. (i) Late (improved) gradient re-writes responses, predicting late specification and stable positions; (ii) Cell sorting corrects wrong positions, predicting early specification and rearrangement afterwards.

Olig2 reporter marks the dorsal boundary of pMN domain in the final pattern (Figure 2.2A). Similar to the *mnx1:gfp*+ cells, *olig2:gfp*+ cells emerge at different positions with negative and positive cells intermingled in a “salt and pepper” fashion, most evidently dorsal-laterally (Figure 2.6D). GFP quantification shows that the olig2 “stripe” arises from a mixed population over a wide range (Figure 2.6E). Nkx2.2a reporter marks the LFP domain and borders *mnx1:gfp* expression ventrally (Figure 2.2A). *Nkx2.2a:mgfp*+ cells start to be detectable around 13hpf at variable locations but become restricted to the stereotypic 2 columns on both sides of the MFP cell after 15hpf (Figure 2.6C). The distribution of LFPs is wide initially and becomes narrower (Figure 2.6D), suggesting the LFP domain also arises from mixed populations, although not as evident as pMNs likely due to the small size (2 cells) of the LFP “stripe”.

Together, these data suggest that specification of ventral progenitors occurs early in spatially mixed distributions, not in sharply segregated “stripes”. These stripes form later in the final pattern. However, as tracking of transgenic marker expression is challenging at earlier times it remains possible that progenitors only transiently express these markers and then either repress

or increase the expression depending on the Shh level at their positions (Dessaud et al, 2010). In this scenario, early heterogeneity in the Shh gradient would be irrelevant as the progenitors remain labile and the early erroneous responses would be overwritten by an improved gradient (Figure 2.5F.i). Alternatively, it is possible that the early response is maintained and these specified cells physically move into the locations that match their specified identity (Figure 2.5F.ii).

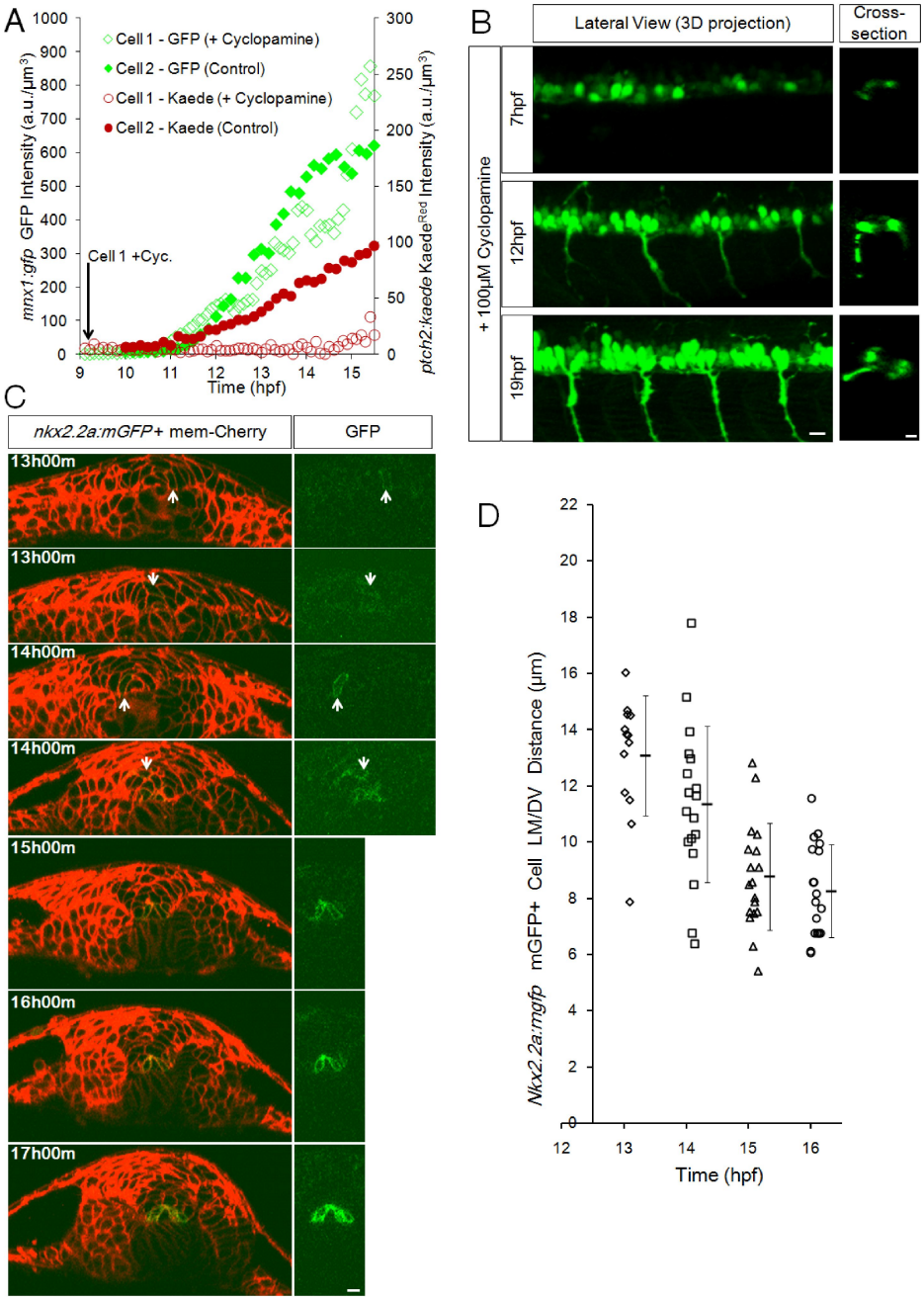


Figure 2.6. Timing of specification and domain formation of pMNs and LFPs (related to Figure 2.5)

(A) Specified cells increase marker expression independently from Shh response. 2 representative tracks from a Cyclopamine treated embryo (Cell 1, Cyclopamine added at the start of imaging near 9hpf) and a control embryo (Cell 2, no Cyclopamine). Embryos are double

Figure 2.6 (*Continued*) transgenic for *ptch2:Kaede* and *mnx1:gfp*. A 405nm laser scan was performed before acquiring each time point to convert Kaede(green) to Kaede(red). 3 additional Cyclopamine tracks and 5 additional control tracks show similar behavior (data not shown).

(B) Representative lateral 3D and cross-sectional views of Cyclopamine treated embryos. Early treatment does not fully remove GFP+ (*mnx1:gfp*) cells, a possible consequence of drug efficiency and/or early pMNs that can be specified in the absence of extracellular Shh (Lewis and Eisen, 2001; Chen et al, 2001). Cross-sections indicate that the specified cells still localize to correct positions. All scale bars: 10μm.

(C) Cross-sectional view of *nkx2.2a:mgfp* expression at different times. Before 13hpf potential mGFP+ cells could not be uncovered by imaging this transgenic line. At 13 and 14hpf GFP+ cells are found in variable locations and the left-right symmetric distribution of GFP+ cells in the LFP domain is not established. After 15hpf most GFP+ cells line up into the LFP domain on both sides of the triangular MFP cell, this position is indicative of LFP fate after 16hpf.

(D) *Nkx2.2a:mgfp*+ cell distribution over time. Each mark represents the location of one mGFP+ cell at the corresponding time point. Note the narrowing of distribution from 14 to 16hpf and the population average (±s.d.) moving closer to the MFP/notochord.

Progenitor divisions are lineage restricted and contribute to cell mixing

To further refine our estimate of the timing of cell specification, we analyzed the lineage trees of identified progenitors. By tracking the mothers and/or grandmothers of specified cells back to as early as neural ectoderm stages we found, strikingly, a strong positive correlation of fate in sister and first cousin cells in the pMN and LFP pools (Figure 2.7A). We did not observe any divisions that generate a pMN and a LFP cell (0/83); more generally the final divisions rarely lead to two

progenitors of different types. Since a great portion of these divisions (30/83) happen before 12hpf (Figure 2.8A), these results argue against the labile cell fate idea, as sister cells adopting different progenitor fates should be found if specification happens late, unless sister cells keep sharing the same position in the Shh gradient (e.g. they are neighbors). To test this, we followed the positional dynamics of sister cells in pairs by measuring their separation distance over time (Figure 2.7B). We found that daughter cells immediately become separated after cell division, even if they become neighbors later, suggesting cell division is one cause of positional mixing. Daughter cells of the same fate can be found on opposite sides of the midline and in different segments of the neural tube (Figure 2.8B), consistent with previous studies (Kimmel et al, 1994; Park et al, 2004). To determine whether these divisions might lead to a difference in sister cell positions in the Shh gradient, we analyzed 50 divisions throughout the LM/DV axis and time (Figure 2.7C, 2.8C). A significant portion (18/50) of divisions happens along the LM/DV axis, so that the positions of the daughter cells relative to the notochord are clearly different. Together, these data show that sister cells share fate but not position at early stages of patterning, suggesting specification (or at least fate bias) has been established in the mother/grandmother cells in a spatially mixed pattern within a dynamic tissue.

Our marker tracking and lineage tracing results show that cells may become specified at "wrong" places due to movements, divisions, and heterogeneous signaling. Additional mechanisms are required to make clean "stripes" from a dynamic, mixed progenitor population. While up to this point cell movement appears to act as a limitation to morphogen patterning precision, could the movements after specification contribute positively to the pattern (Figure 2.5F.ii)?

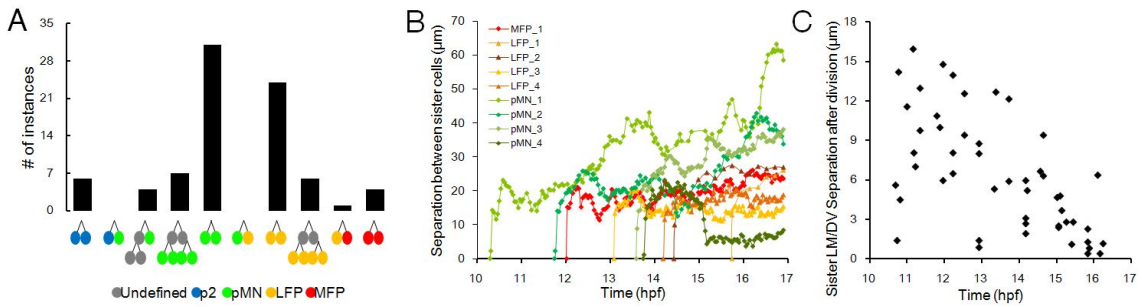


Figure 2.7. Progenitors share fate but not position with sisters and cousins at early stages

(A) Summary of lineage motif counts (n=83). Counts are collected from 18 independent datasets. Motifs with 2 generations are not often captured in the imaged time window so the count does not suggest that 2-generation motifs happen in lower frequencies than 1-generation motifs. Division times: before 12hpf, n=30, 12-14hpf, n=20, after 14hpf, n=33. See also Figure 2.8A.

(B) Separation dynamics of sister cells after birth. The 0 points: the birth time of sister cells from the division of mother cell. A distance of 6 to 8μm indicates the sisters remain neighbors, 10 to 16μm one cell separation, etc. See also Figure 2.8B.

(C) Cell divisions causing position instability. 50 division events randomly picked through time. 18 divisions happen closely along the LM/DV axis, generating at least one cell-diameter difference (>8μm) in position between sister cells. At later time, more divisions are perpendicular to the LM/DV axis, generating no significant positional difference between sisters (<3μm). See also Figure 2.8C.

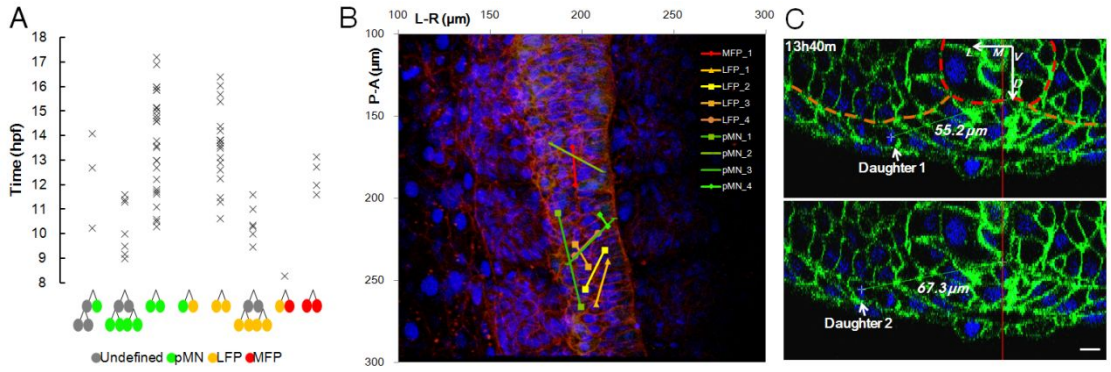


Figure 2.8. Positional separation and mixing of sister cells (related to Figure 2.7)

(A) Distribution of divisions in time by division types. 66 out of 83 divisions from Figure 2.7A are plotted (accurate times of some divisions were not recorded thus were not included). Divisions giving rise to pMNs and LFPs spread out along the same time window but a division giving rise to one pMN and one LFP was not found.

(B) Separation of sister cells. Image is a 3D projection of the original dataset at 16.9hpf, lines connect positions of sister cells. Due to different Z depth, the cells under the marks may not be the actual tracked cells, which might lie deeper below. A lot of sister cell pairs are separated by the midline, pMN_1 sister pair is separated into different neural segments.

(C) Example of divisions that render one daughter closer to notochord than the other. Images are cross-section slices rendered and measured in GoFigure2. Red line: midline. Scale bar: 10μm.

Cell sorting establishes sharply bordered progenitor domains

To understand how cell movements contribute to patterning, we tracked cell type identified cells back in time. Surprisingly, we found that the early distribution of the progenitors does not match their final distribution in terms of relative position or order (Figure 2.9A,B). In this fully tracked ventral neural segment, progenitors that make the pattern (Figure 2.9A.ii,iv) are initially spread out and mixed with cells that will not join this segment (Figure 2.9A.i,iii). Moreover, future

pMNs may start off either touching the notochord or located far away from the notochord. Later, all these pMNs come together to locate into a sharply bordered domain (Figure 2.9A). LFPs and MFPs also show similar behavior, albeit in smaller spatial ranges compared to pMNs (Figure 2.9B, 2.10A). The early distributions of tracked MFPs, LFPs and pMNs resemble the wide and mixed patterns of early *shh:gfp*, *nkx2.2a:mgfp* and *mnx1:gfp* expression, respectively (Figure 2.3A, 2.6C, 2.5A). While we found that the MFPs always touch the notochord and line up along the midline earliest (Figure 2.9B,C), pMNs and LFPs frequently intermingle and switch positions (Figure 2.9B,D, Movie S2.5). These rearrangements happen most often as cells enter the neural keel and after divisions. For example, in Movie S2.5, at 11hpf a LFP progenitor was initially located more lateral to a pMN in the neural plate. The LFP progenitor migrated dorsal to the pMN at around 12hpf and remained no closer to the notochord than the pMN. It divided around 13hpf generating 2 future LFPs. As a result of this division, one daughter LFP was further dorsal compared to the pMN until around 14hpf when it moved to equal distances. Finally, after 14.5hpf this daughter LFP inserted between an MFP and the pMN and maintained that position onwards. By locating the tracked cells in fully segmented neural plate/tube at different times (Figure 2.9E), we found that the pMN/LFP boundary marked by *mnx1:gfp* expression starts to emerge between tracked cells after 14.5hpf; at times earlier than this the pMNs and LFPs are located in wide ranges that overlap. Most ventral cells settle into stable positions by 15hpf (Figure 2.9B). We verified cell fates by tracking with additional fate markers, and determined that these cells stay stably within their domains by later stage movies (Figure 2.10B-D, Movie S2.6, data not shown). These data demonstrate that cell sorting directly establishes the “French Flag” pattern. The fact that cells at initially widely separated locations can have the same fate and final location whereas initial neighbors may have different final locations and fates is unexpected. This observation is

inconsistent with the positional specification model. However, we note that on the population level a rough correlation between position and fate exists throughout, and is sharpened over time by cell sorting (Figure 2.9B, insert). Together, our data rule out the notion that naïve cells are specified between spatial thresholds and remain in the same relative positions; instead, the progenitor domains and their boundaries form by sorting of specified cells from widely dispersed locations.

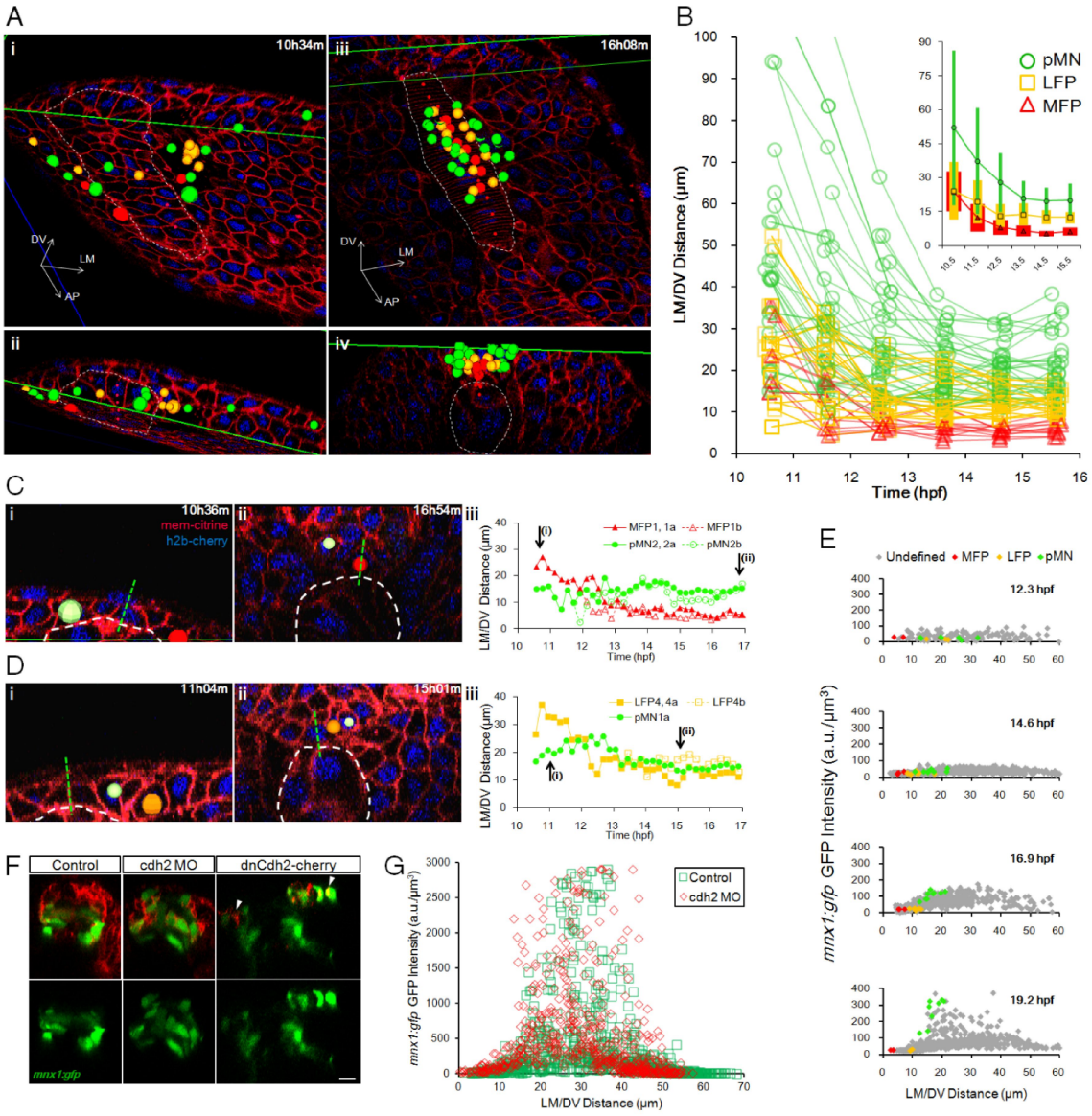


Figure 2.9. Progenitors enter stable locations and form sharp boundaries by intensive cell

Figure 2.9 (*Continued*) **rearrangement**

(A) Distribution of tracked cells from a fully analyzed ventral neural segment (comprised of 7 MFPs, 13 LFPs and >20 pMNs) at early neural plate stage (i,iii) and neural tube stage (ii,iv). (ii), (iv): Corresponding cross-sectional views of (i), (iii). Green lines indicate the intersection of cross-section view and dorsal view (i, iii and lower line in ii) or the upper boundary of the dataset (iv and upper line in ii). Colored spheres: 3D locations of tracked cells (Red: MFP; Orange: LFP; Green: pMN). Dashed lines: notochord boundary. Small red spheres: notochord top midline.

(B) Trajectories of tracked cells along the LM/DV axis demonstrating intensive sorting. For simplicity, only six time points on the tracks are plotted. 66 tracks collected from 4 datasets are plotted. Some cells exhibit rearrangements beyond 16hpf. Insert: population average position ±s.d. (colored bars) of tracks by cell type plotted on the same axes. See also Figure 2.10A.

(C) Example of relative positional changes of a pMN (light green, pMN2a) and a MFP (red, MFP1,1a). Green dashed line: midline. White dashed line: notochord boundary. (iii): Full movement trajectories of the cells (same in D, for simplicity, one of the daughter cell tracks is continued with the mother track).

(D) Example of positional switch between a pMN (light green, pMN1a) and a LFP cell (orange, LFP4,4a). See also Movie S2.5.

(E) *mnx1:gfp* expression boundary formation between LFPs and pMNs. GFP intensity distribution by position plotted for 4 time points. Each mark represents a cell (>200 cells per time point). Colored marks: tracked cells with known fates; Grey marks: other segmented cells at the plotted time point.

(F) Cdh2 perturbations on *mnx1:gfp*+ domain formation. Images are 24hpf cross-sections of

Figure 2.9 (*Continued*) mosaic labeled (cherry ±cdh2 Morpholino (MO) and dominant negative Cdh2-cherry fusion (dnCdh2-cherry)) neural tubes. Arrowheads: puncta of dnCdh2-cherry. Scale bar: 10μm. See also Figure 2.10E, Movie S2.7.

(G) Quantification of GFP+ cell distribution in Cdh2 morphant and control. See also Figure 2.10E.

Our results suggest that cell sorting is required for pattern formation in the neural tube. To test this hypothesis, we mosaically perturbed cadherin-2 (cdh2), a neural adhesion molecule expressed by all neural progenitors and required for their movements, using a cdh2 morpholino and a dominant negative version of cdh2 (Lele et al, 2002; Rieger et al, 2009). In the perturbed embryos, many *mnx1:gfp*+ cells are misplaced in a wider and more mixed pattern at stages by which "stripes" have formed in controls (Figure 2.9F,G, 2.10E). Live imaging of perturbed cells reveals that their misplacement resulted from reduced integration into the neural keel/tube, which likely blocked cell sorting and thus preserved the noisy spatial pattern of specification (Movie S2.7). These data indicate proper adhesion is required for cell sorting and consequently pattern formation.

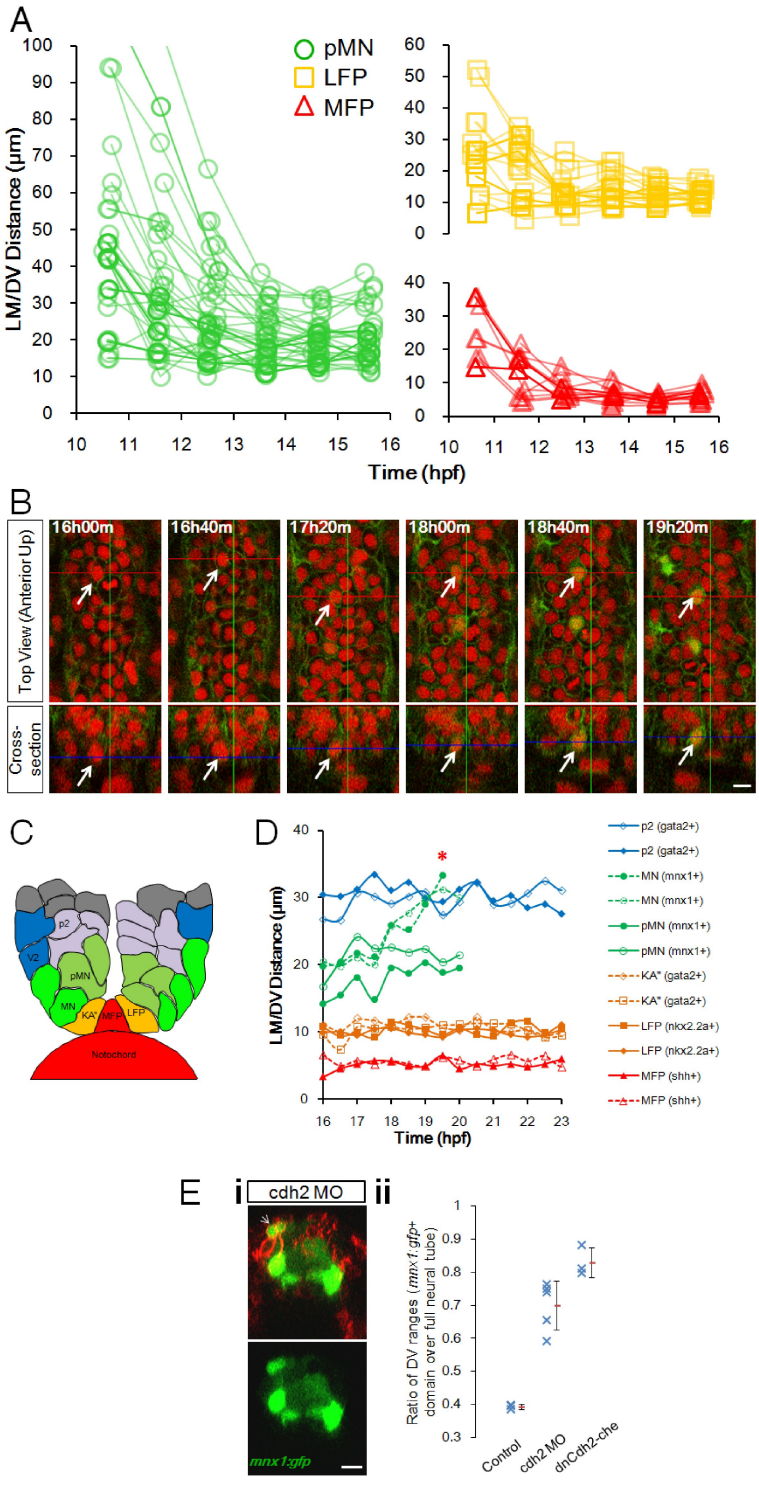


Figure 2.10. Domain boundaries remain stable after 16hpf, cell type verifications by additional markers and additional cdh2 perturbation data (related to Figure 2.9)

(A) Plots by cell type from Figure 2.9B.

Figure 2.10 (*Continued*) (B) LFP cell type confirmation by later stage timelapse of *tg(gata2:gfp)* embryos. KA" neurons (*gata2:gfp*+) are born from the LFP domain among LFP progenitors (*nkx2.2a:mgfp*+, Figure 2.2A, 2.6C; Huang et al, 2012). Note that the tracked LFP cell stays in stable location over time and becomes gata2:gfp+ by 18h40m. All scale bars: 10μm.

(C) Schematic illustration of cell positions in the final pattern. Enlarged and annotated from Figure 2.1Di.

(D) Later stage tracking with various markers. Asterisk: Maturing MNs leave pMN domain and migrate laterally, effectively increasing the LM/DV distance. Neuron movements happen after progenitor domain formation and are distinct from progenitor movements. Cell positions are otherwise stable in this time window indicating stability of established pattern. See also Movie S2.6.

(E) (i).Cdh2 perturbed embryo as in Figure 2.9F. Example of ectopically localized pMNs that are cherry+ (arrow). Both injected and uninjected (see also Figure 2.9F) pMNs might become misplaced, suggesting the pattern disruption caused by Cdh2 MO is cell non-autonomous. (ii). Cdh2 perturbed embryos have wider distribution of pMNs. The ratio between the maximum DV range of the *mnx1:gfp*+ cell cross-sectional distribution over the full length of the neural tube at 24hpf was measured. dnCdh2-che: dominant negative Cdh2-cherry fusion. Blue marks: measurement on one embryo. Red marks: average ratios (±s.d.).

Ectopically induced pMNs migrate to form a sharp domain

A model in which pattern forms by sorting of specified cells predicts that ectopically induced progenitors should migrate to the correct positions corresponding to their fates. To test this prediction, we mosaically over-expressed the transcription factor Mnx2a, which is a marker of

pMN (Wendik et al, 2004) and whose homolog MNR2 induces ectopic MNs in chick embryos (Tanabe et al, 1998). We injected *mnx2a* mRNA mixed with *mCherry* mRNA into one blastomere at the 8-16 cell stage. Strikingly, we found strong ventral segregation of mCherry labeled cells to the normal pMN domain in *mnx2a* injected embryos whereas control embryos show a random distribution of labeled cells across the DV axis (Figure 2.11A,B). In class I embryos, the pMN domain is fully occupied by descendants of the injected blastomere, and motor axons are strongly labeled evenly along the body axis (Figure 2.12A), a phenomenon never observed in control injections. Conversely, in the dorsal domains of the neural tube, in contrast to control embryos, injected cells are missing in *mnx2a* injected embryos (Figure 2.12A,B). Mosaic overexpression of Mnx1 resulted in a similar phenotype (Figure 2.11B, data not shown). Quantification of GFP+ MNs further confirms that Mnx2a injected cells contribute more extensively to the pMN domain as compared to random contribution of control *mCherry* injected cells (Figure 2.11D). Cyclopamine treatment of injected embryos starting at 7hpf does not alter the Mnx2a phenotypes (Figure 2.11B, 2.12C), suggesting that the ventral localization of Mnx2a injected cells is independent of Shh response. Indeed, early specified normal pMNs also form a smaller but sharp domain in the presence of Cyclopamine (Figure 2.6B). Mnx2a injection alleviates reduction of MNs by Cyclopamine treatment, and the injected cells express pMN markers and maintain progenitor location and morphology (Figure 2.11D, 2.12D), confirming that they have become specified to pMNs. To understand how the Mnx2a phenotypes arise at the cellular level, we tracked the movement of injected cells. Interestingly, these cells form clusters in the early neural plate (Figure 2.12E, data not shown) suggesting adhesion changes accompanying specification. They migrate together to populate the ventral domains (Figure 2.11C) to give rise to Mnx2a phenotypes, and their ventral bias becomes evident after

intercalating into the neural keel (Figure 2.11C, compare 14.5 and 16.5hpf), similar to normal pMNs. These data suggest that Mnx2a may control specific adhesion affinities of pMNs that control their sorting. Indeed, in cdh2 morpholino and Mnx2a co-injected embryos, despite severe disruption of morphogenesis, the injected cells remain co-localized and ventrally biased (Figure 2.11B, 2.12F).

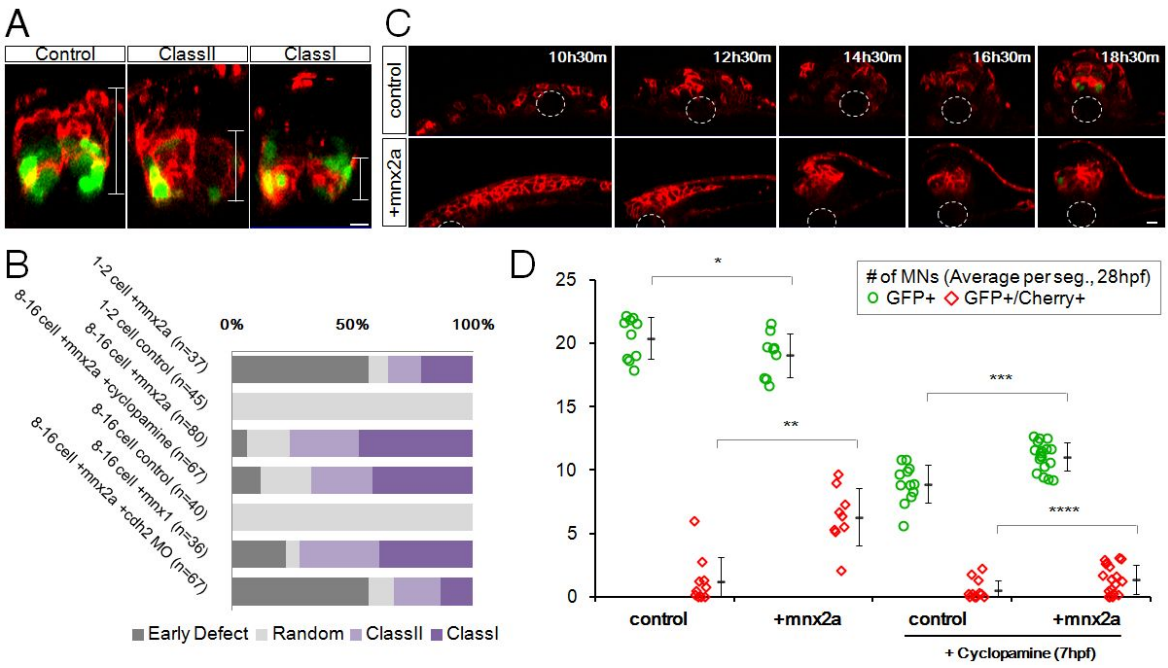


Figure 2.11. Ectopic Mnx2a expressing cells form a sharp ventral domain similar to the pMN domain

(A) 24hpf neural tube phenotypes after injection of *mem-mCherry* ±*mnx2a* mRNAs in one blastomere at 8-16 cell stage. Phenotypes are classified according to the distribution of mCherry+ cells (Brackets): class I embryos contain cells only in the ventral 1/3 of the neural tube; class II embryos contain cells in the ventral 2/3; "random" contains injected cells throughout. Green: *mnx1:gfp*. All scale bars: 10µm. See also Figure 2.12A.

(B) Summary of mosaic injection experiments. Early defect embryos failed to form neurula. Cyclopamine treatment started at 7hpf.

Figure 2.11 (*Continued*) (C) Sample time course of Mnx2a domain formation. This Mnx2a embryo became class II type. Dashed line circles: position of the notochord. Green: *mnx1:gfp*. Red: mem-mCherry. See also Figure 2.12B.

(D) Mnx2a expressing cells replace "normal" pMNs. Imaging and counting of MNs as Figure 2.5C. p values: *0.09;**0.00004,***0.0001,****0.03 (Student's t test).

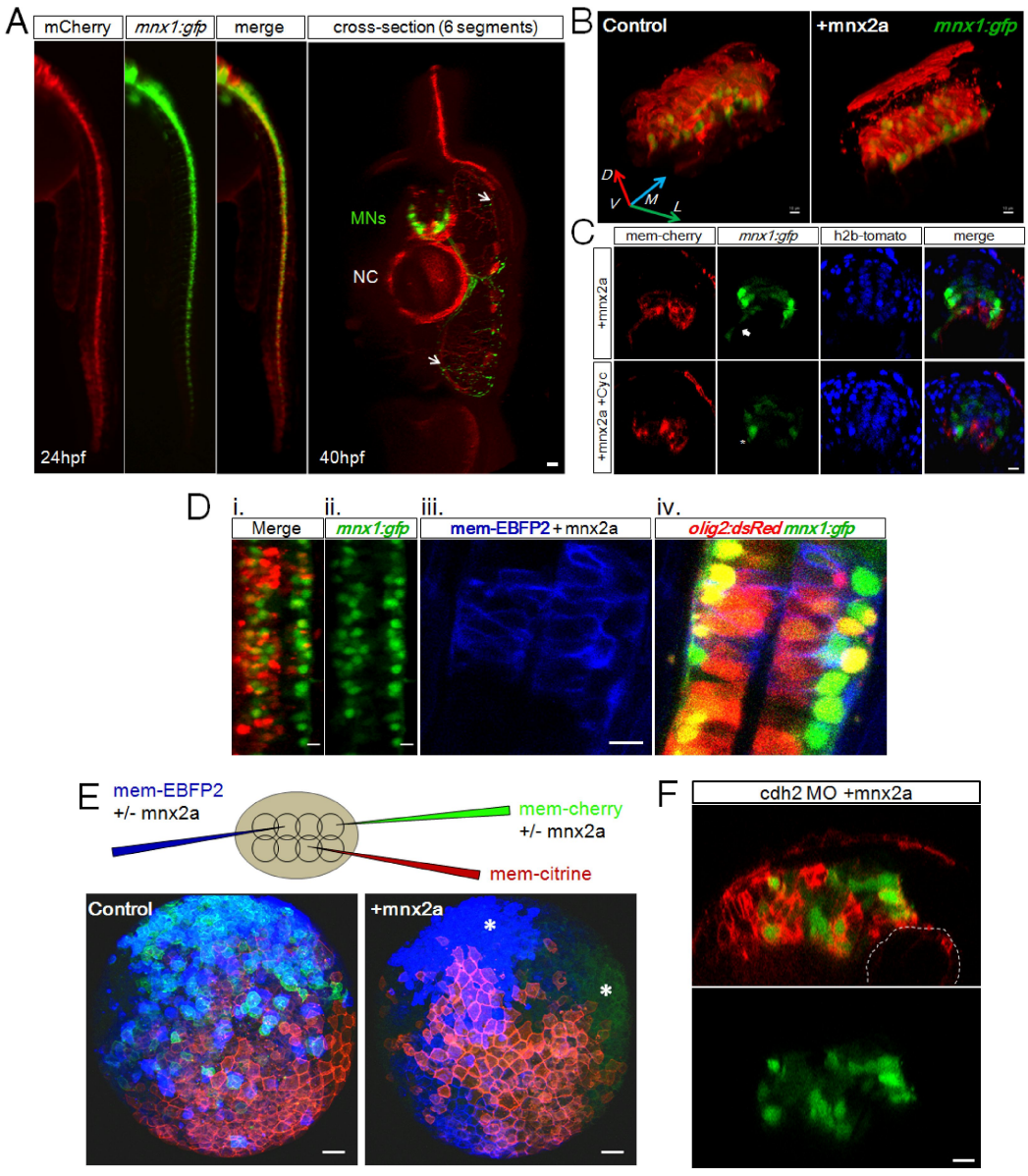


Figure 2.12. Ectopic Mnx2a expressing cells populate pMN domain in a Shh independent

Figure 2.12 (*Continued*) **manner, Mnx2a injected cells exhibit changed adhesion properties (related to Figure 2.11)**

(A) Mnx2a injected cells localize to pMN domains and become MNs. Left: Lateral view of a ClassI embryo, the Mnx2a+mCherry co-injected cells predominantly segregated to the ventral neural tube, overlapping with the transgenic expression domain of *mnx1:gfp*. Right: mCherry+ cells become MNs. Arrows: mCherry+ motor axons in the muscle tissue. All scale bars (except E) are 10μm.

(B) Mnx2a cells populate extensively and exclusively the ventral neural tube. Images are 3D projections of original 2-channel z-stacks of 24hpf embryos. Unlike control embryos whose injected cells are randomly scattered in the neural tube (left), mnx2a injected cells form a clear ventral domain (right). The mCherry+ patch of cells seen above the neural tube are skin cells.

(C) Cyclopamine treatment does not alter the Mnx2a phenotypes. 100μM Cyclopamine was applied to dechorionated *tg(mnx1:gfp/actb2:h2b-tdTomato)* embryos at 7hpf. The axonal projections of the MNs were affected (Arrow, asterisk, Charron et al, 2003) but the pMN domain "stripe" formed by Cherry+ cells was not affected.

(D) Mnx2a injected cells become pMNs. (i, ii). Top view of example *tg(mnx1:gfp)* embryo injected with Mnx2a and H2B-mCherry mRNA and treated with Cyclopamine at 12hpf, whose left half is populated by mCherry+ cells and have increased number of GFP+ cells compared to the right half. (iii, iv). Top view of example *tg(mnx1:gfp/olig2:dsRed)* embryo injected with Mnx2a and mem-EBFP2 mRNA. EBFP2+ cells are dsRed+ and show progenitor location and morphology.

(E) Mnx2a injected cells form clusters at early neural plate stage. 3D projections (animal/dorsal view) of typical 8hpf triple injected embryos. Injections are performed as shown in the

Figure 2.12 (*Continued*) illustration. The formation of large clusters by Mnx2a cells as compared to random mixing of control cells suggests changes of adhesion properties in these cells, which may play a role in cell sorting. The injection amount of Mnx2a mRNA is higher (20ng/µl) in these experiments compared to Figure 2.11. Scale bar: 50µm.

(F) Mnx2a and cdh2 MO co-injected cells still form clusters and localize ventrally despite strong morphogenesis defects. Injections are performed as in Figure 2.11A. Many embryos show severe early defects and the survivors often show distorted neural tubes (image is a typical example). Mnx2a induced adhesion still appears to be in effect in these embryos. Dashed lines: Notochord boundary.

Our results show that induced ectopic progenitors move to form sharp domains similar to the normal pattern in a Shh signaling independent manner. They further suggest that specification creates adhesive differences between cells of different fates. Together, these data support a model in which specified progenitors self-assemble into precise spatial domains by cell adhesion dependent cell sorting.

DISCUSSION

The role of cell movement in neural plate/tube patterning

We captured the first 4-dimensional picture of pattern formation from early neural plate to neural tube at single-cell, trackable resolution in zebrafish, and discovered that the sharply delineated pattern of progenitor domains forms through sorting of specified cells. Our observations

challenge and extend the classic positional specification model in several ways. First, the classic model assumes graded response as a function of position in a smooth, monotonic morphogen gradient. While we do not know how closely the distribution of Shh molecules in the neural plate resembles such a gradient, we have shown that the response is highly dynamic and heterogeneous. We suggest that, even if Shh morphogen forms a perfect gradient, the movements of the cells will inevitably complicate their Shh exposure, making the response pattern noisy. Second, the classic model suggests naïve cells become specified at stereotypic positions. We have shown fate markers are expressed in intermingled patterns during and preceding more cell movements. In addition, cell fates are lineage restricted (e.g. pMN vs. LFP) long before the final pattern emerges. These observations indicate that specification and positioning are separate in time and cells get specified outside stereotypic positions. Third, the classic model interprets the ventral to dorsal progressing pattern of Shh mediated gene expression (Jeong and MacMahon, 2005) as a result of stationary cells changing their gene expression as they accumulate more Shh signals (Chamberlain et al, 2008; Dessaud et al, 2010). Our observations suggest that, in addition to gene expression changes, cells can maintain their gene expression state and physically move to contribute to the refining pattern (e.g. a pMN moves away from the notochord). Together, we propose a revised model for neural tube patterning incorporating imprecision of positional information and cell movement (Figure 2.13, 2.14): Cell positions are unstable in the dynamic tissue and morphogen signaling across the tissue is spatially noisy (at least in part due to movement of responding cells), resulting in a “salt and pepper” specification pattern. Cell sorting then segregates different progenitors into sharply defined domains.

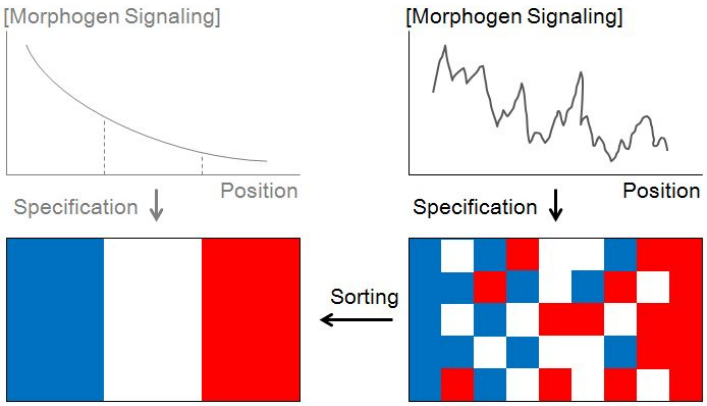


Figure 2.13. Revised "French Flag" model incorporating dynamics of morphogen gradient and cell sorting

This model depicts specification and sorting sequentially for conceptual clarity but they occur at different and overlapping times for different cells. See Discussion. See also Figure 2.14.

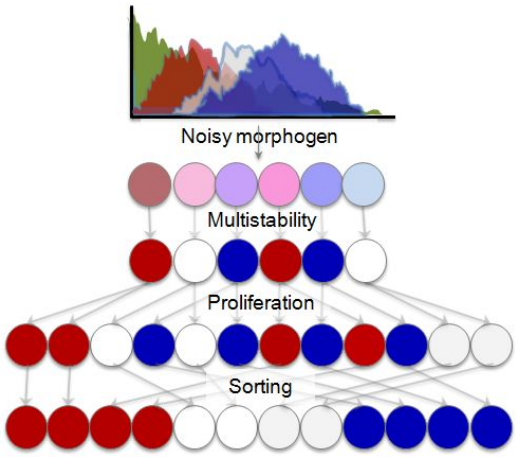


Figure 2.14. Molecular and cellular mechanisms orchestrate "French Flag" patterning (related to Figure 2.13)

Model of neural tube pattern formation. Specification by an evolving morphogen gradient (green) is noisy and imprecise spatially. An intracellular gene regulatory network converts the graded Shh input into discrete cell fates in a process called multistability (Balaskas et al, 2012), but in a positionally mixed manner. Cell divisions amplify the mixed progenitor pool, and specified cells

Figure 2.14 (*Continued*) move to build the final pattern. We have illustrated these processes sequentially for clarity but they overlap in time. See Discussion.

While important for neural tube patterning in zebrafish, the role of cell movement in other vertebrates such as chick and mice remains to be elucidated and may be different or context-dependent. The modes of neural tube morphogenesis among vertebrates vary considerably presumably depending on the degree of progenitor epithelialization (Smith and Schoenwolf 1997; Clarke 2009). For example, in primary neurulation that occurs in the anterior neural tubes of chick and mice, an epithelialized cell sheet undergoes a folding process that forms a lumen through invagination (Smith and Schoenwolf 1997). In secondary neurulation that occurs more posteriorly, however, a neural rod of less epithelialized cells forms first which then cavitates to form a lumen *de novo* as cells epithelialize (Catala et al, 1996). We have observed that more epithelialized cells have less mobility in zebrafish. The higher degree of epithelialization in primary neurulation suggests that there is unlikely as much cell mixing or rearrangement as in the zebrafish neural tube, which shares more similarities with secondary neurulation (Clarke 2009). The amount of cell movement in chick and mice neural tubes has been assessed by clonal labeling studies (Leber et al, 1995; Inoue et al, 2000; Das and Storey, 2012). These studies show that there is wide cell dispersion at early stages but little cell movement later. Unfortunately, the exact trajectories of these cell movements, times of cell divisions and how they relate to Shh responses and specification are not clear. To determine to what extent (if any) cell sorting contributes to neural tube patterning in these vertebrates, imaging data comparable to ours in spatial temporal resolution and coverage is required. We note that ongoing efforts towards these goals show promising potential (Yamaguchi et al, 2011; Das and Storey, 2012).

The mechanisms controlling cell sorting

We have not yet determined the molecular details of cell sorting, but our data suggest it is a complex process likely orchestrated by multiple adhesion molecules. We have shown Cdh2 is required for proper pMN domain formation. In addition, Mnx2a appears to cause adhesion changes that drive sorting of ectopic pMNs, suggesting specification downstream of Shh signaling may activate fate specific affinities, as observed in the abdomen of Drosophila (Lawrence et al, 1999). In our movies we also found cell rearrangements happen most often as cells mix during intercalation while forming the neural keel and after divisions, conditions that likely facilitate the effect of short-ranged adhesion forces. Disruption of such intercalation results in misplaced progenitors. Previously it has been shown that differential adhesion can mediate migration and pool sorting of post-mitotic neurons (Price et al, 2002), a process that follows progenitor domain formation. A similar strategy might be employed by the progenitors, as they also express different cadherins and protocadherins in conserved patterns along the DV axis (Lin et al, 2012). What sets of specific adhesion molecules correspond to different progenitor fates and how they are regulated and cooperate to control cell sorting remains to be elucidated.

An alternative sorting mechanism is chemotaxis of specified cells, in which the direction and final location of sorting is determined by diffusible signals while adhesion molecules only serve as the structural necessity for cells to move. We have shown Shh response is not required for sorting but it remains possible non-canonical Shh or other molecular gradients (e.g. Bmp, Wnt) provide positional cues for cell movement.

Cell self-assembly confers robustness to positional noise and errors

The formation of spatially distinct domains faces noise at multiple scales, including molecular noise as described previously (Paulsson, 2004; Lander et al, 2009) and cell positional noise caused by stochastic cell movements as described here. We believe that multiple strategies are used to achieve robust patterning in the face of this noise. The intracellular GRN (Balaskas et al, 2012) can help make and maintain correct fate decisions by canalizing noisy signaling inputs into discrete, non-overlapping states of gene expression and thus cell fates. Intercellular interactions, such as cell sorting shown in this study, allow overlapping distributions of cell types caused by spatially noisy signaling to be corrected. Furthermore, other intercellular interactions such as lateral inhibition may play a role in size control of progenitor domains. For example, in Mnx2a injected embryos the final number of pMNs seems to be regulated despite being initially too large, suggesting ectopic Mnx2a expressing cells may prevent uninjected cells from becoming pMNs. Characterizing the molecular and cellular details of these different interactions will be vitally important for understanding how embryos canalize molecular and positional noise as well as genetic and environmental variation to attain developmental norms (Waddington 1942).

In summary, our study highlights the power and importance of live observation of cell behavior in understanding developmental patterning, and provides a model of how patterns robustly arise in the dynamic environment of the developing ventral neural tube. Cell sorting by differential affinities is a classical idea (Sternberg, 1963) alongside the morphogen model (Wolpert, 1969). There is no reason to think that Shh signaling is unique in showing a highly dynamic, noisy pattern of activity. If these are general features of morphogens, then self assembly may be a general mechanism to assign positions to specified cells, as cell movement is common during the morphogenesis and proliferation of both signaling centers and their target

fields (Kay and Thompson, 2009).

EXPERIMENTAL PROCEDURES

Zebrafish strains and maintenance

See Extended Experimental Procedures for protocols, sources and references for transgenic strains used in this study. All fish related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Microinjections of mRNAs

For *in toto* labeling, 1-cell stage embryos were injected (Nanoject) 2.3nl 40ng/µl of labeling mRNA(s) (h2b-cherry, mem-citrine, mem-cherry, mem-EBFP2 and combinations). For mosaic injections, one blastomere of 8 to 32-cell stage embryos was injected with approximately 1nl 20ng/µl one labeling mRNA with or without 10ng/µl *mnx2a, dnCdh2-cherry* mRNA.

Timelapse 2-photon/confocal imaging

Live imaging was performed using a Zeiss 710 confocal/2-photon microscope (objective: C-Apochromat 40X 1.2 NA) with a home-made heating chamber maintaining 28ºC. Chameleon (Coherent) laser line 1020nm was used for 3-channel 2-photon *in toto* sessions. See Figure 2.2, Extended Experimental Procedures for details.

Image Data analysis

Nuclear segmentation and tracking was performed using GoFigure2, an open-source, cross-platform software application we have developed for image analysis (www.gofigure2.org). Segmentation and track tables exported from GoFigure2 were further processed and plotted with Microsoft Excel. See Figure 2.2, 2.4, Extended Experimental Procedures for details.

ACKNOWLEDGEMENTS

We thank D. D'India for fish care, H. Otsuna for movie making, R. Koster, D. Meyer, U. Strähle, B. Appel, S. Higashijima, S. Lin and L. Zon for sharing reagents. D. Tulga, T. Hiscock for technical assistance, Q. Mao, R. Noche, A. Green, Y. Chen, W. Ma, R. Ward, A. McMahon, M. Kirschner, C. Cepko and T. Mitchison for comments. This work is supported by NIH grants HG004071 and DC010791. F. X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

REFERENCES

Ahn, S., and Joyner, A.L. (2004). Dynamic changes in the response of cells to positive hedgehog signaling during mouse limb patterning. Cell *118*, 505-516.

Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K.M., Briscoe, J., and Ribes, V. (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell *148*, 273-284.

Catala, M., Teillet, M.A., De Robertis, E.M., and Le Douarin, M.L. (1996). A spinal cord fate map in the avian embryo: while regressing, Hensen's node lays down the notochord and floor plate thus joining the spinal cord lateral walls. Development *122*, 2599-2610.

Chamberlain, C.E., Jeong, J., Guo, C., Allen, B.L., and McMahon, A.P. (2008). Notochord-derived Shh concentrates in close association with the apically positioned basal body in neural target cells and forms a dynamic gradient during neural patterning. Development *135*, 1097-1106.

Clarke, J. (2009). Role of polarized cell divisions in zebrafish neural tube formation. Curr. Opin. Neurobiol. *19*, 134-138.

Ciruna, B., Jenny, A., Lee, D., Mlodzik, M., and Schier, A.F. (2006). Planar cell polarity signalling couples cell division and morphogenesis during neurulation. Nature *439,* 220-224.

Das, R.M., and Storey, K.G. (2012). Mitotic spindle orientation can direct cell fate and bias Notch activity in chick neural tube. EMBO Rep. *13,* 1030.

Dessaud, E., Yang, L.L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitch, B.G., and Briscoe, J. (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature *450,* 717-720.

Dessaud, E., Ribes, V., Balaskas, N., Yang, L.L., Pierani, A., Kicheva, A., Novitch, B.G., Briscoe, J., and Sasai, N. (2010). Dynamic assignment and maintenance of positional identity in the ventral neural tube by the morphogen sonic hedgehog. PLoS Biol. *8,* e1000382.

Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell *75,* 1417-1430.

Ericson, J., Briscoe, J., Rashbass, P., van Heyningen, V., and Jessell, T.M. (1997). Graded sonic hedgehog signaling and the specification of cell fate in the ventral neural tube. Cold Spring Harb. Symp. Quant. Biol. *62,* 451-466.

Erskine, L., Patel, K., and Clarke, J.D. (1998). Progenitor dispersal and the origin of early neuronal phenotypes in the chick embryo spinal cord. Dev. Biol. *199,* 26-41.

Flanagan-Steet, H., Fox, M.A., Meyer, D., and Sanes, J.R. (2005). Neuromuscular synapses can form in vivo by incorporation of initially aneural postsynaptic specializations. Development *132,* 4471-4481.

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P., and Tabin, C.J. (2004). Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell *118,* 517-528.

Huang, P., Xiong, F., Megason, S.G., Schier, A.F. (2012). Attenuation of Notch and Hedgehog signaling is required for fate specification in the spinal cord. PLoS Gen *8(6),* e1002762.

Inoue, T., Nakamura, S., and Osumi, N. (2000). Fate mapping of the mouse prosencephalic neural plate. Dev. Biol. *219,* 373-383.

Jeong, J., and McMahon, A.P. (2005). Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1. Development *132,* 143-154.

Jessen, J.R., Meng, A., McFarlane, R.J., Paw, B.H., Zon, L.I., Smith, G.R., and Lin, S. (1998). Modification of bacterial artificial chromosomes through chi-stimulated homologous recombination and its application in zebrafish transgenesis. Proc Natl Acad Sci U S A *95,* 5121-5126.

Jessell, T.M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat. Rev. Genet. *1,* 20-29.

Kay, R.R., and Thompson, C.R. (2009). Forming patterns in development without morphogen gradients: scattered differentiation and sorting out. Cold Spring Harb Perspect. Biol. *1,* a001503.

Krauss, S., Concordet, J.P., and Ingham, P.W. (1993). A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. Cell *75,* 1431-1444.

Kimmel, C.B., Warga, R.M., and Kane, D.A. (1994). Cell cycles and clonal strings during formation of the zebrafish central nervous system. Development *120,* 265-276.

Korzh, V., Edlund, T., and Thor, S. (1993). Zebrafish primary neurons initiate expression of the LIM homeodomain protein Isl-1 at the end of gastrulation. Development *118,* 417-425.

Lander, A.D., Lo, W.C., Nie, Q., and Wan, F.Y. (2009). The measure of success: constraints, objectives, and tradeoffs in morphogen-mediated patterning. Cold Spring Harb Perspect. Biol. *1,* a002022.

Lawrence, P.A., Casal, J., and Struhl, G. (1999). The hedgehog morphogen and gradients of cell affinity in the abdomen of Drosophila. Development *126,* 2441-2449.

Leber, S.M., and Sanes, J.R. (1995). Migratory paths of neurons and glia in the embryonic chick spinal cord. J. Neurosci. *15,* 1236-1248.

Lek, M., Dias, J.M., Marklund, U., Uhde, C.W., Kurdija, S., Lei, Q., Sussel, L., Rubenstein, J.L., Matise, M.P., Arnold, H.H., Jessell, T.M., and Ericson, J. (2010). A homeodomain feedback circuit underlies step-function interpretation of a Shh morphogen gradient during ventral neural patterning. Development *137,* 4051-4060.

Lele, Z., Folchert, A., Concha, M., Rauch, G.J., Geisler, R., Rosa, F., Wilson, S.W., Hammerschmidt, M., and Bally-Cuif, L. (2002). Parachute/n-Cadherin is Required for Morphogenesis and Maintained Integrity of the Zebrafish Neural Tube. Development *129,* 3281-3294.

Lewis, K.E., and Eisen, J.S. (2003). From cells to circuits: development of the zebrafish spinal cord. Prog. Neurobiol. *69,* 419-449.

Lin, J., Wang, C., and Redies, C. (2012). Expression of delta-protocadherins in the spinal cord of the chicken embryo. J. Comp. Neurol. *520,* 1509-1531.

Marti, E., Takada, R., Bumcrot, D.A., Sasaki, H., and McMahon, A.P. (1995). Distribution of Sonic hedgehog peptides in the developing chick and mouse embryo. Development *121,* 2537-2547.

Megason, S.G., and Fraser, S.E. (2003). Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. Mech. Dev. *120,* 1407-1420.

Mosaliganti, K.R., Noche, R.R., Xiong, F., Swinburne, I.A., Megason, S.G. (2012). ACME: Automated Cell Morphology Extractor for Comprehensive Reconstruction of Cell Membranes. PLoS Comp. Biol. *8(12)*: e1002780.

Nicol, A., Rappel, W., Levine, H., and Loomis, W.F. (1999). Cell-sorting in aggregates of Dictyostelium discoideum. J. Cell. Sci. *112 (Pt 22),* 3923-3929.

Park, H.C., Shin, J., and Appel, B. (2004). Spatial and temporal regulation of ventral spinal cord precursor specification by Hedgehog signaling. Development *131*, 5959-5969.

Paulsson J. (2004). Summing up the noise in gene networks. Nature. *427*, 415-8.

Price, S.R., De Marco Garcia, N.V., Ranscht, B., and Jessell, T.M. (2002). Regulation of motor neuron pool sorting by differential expression of type II cadherins. Cell *109*, 205-216.

Rieger, S., Senghaas, N., Walch, A., and Koster, R.W. (2009). Cadherin-2 controls directional chain migration of cerebellar granule neurons. PLoS Biol. *7,* e1000240.

Schafer, M., Kinzel, D., Neuner, C., Schartl, M., Volff, J.N., and Winkler, C. (2005). Hedgehog and retinoid signaling confines nkx2.2b expression to the lateral floor plate of the zebrafish trunk. Mech. Dev. *122,* 43-56.

Schoenwolf, G.C. (1991). Cell movements driving neurulation in avian embryos. Development *Suppl 2*, 157-168.

Shin, J., Park, H.-C., Topczewska, J.M., Mawdsley, D.J., and Appel, B. (2003) Neural cell fate analysis in zebrafish using olig2 BAC transgenics. Meth. Cell Sci. *25(1-2)*: 7-14.

Shkumatava, A., Fischer, S., Muller, F., Strahle, U., and Neumann, C.J. (2004). Sonic hedgehog, secreted by amacrine cells, acts as a short-range signal to direct differentiation and lamination in the zebrafish retina. Development *131,* 3849-3858.

Smith, J.L., and Schoenwolf, G.C. (1997). Neurulation: coming to closure. Trends Neurosci. *20,* 510-517.

Steinberg, M.S. (1963). Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. Science *141,* 401-408.

Surmeli, G., Akay, T., Ippolito, G.C., Tucker, P.W., and Jessell, T.M. (2011). Patterns of spinal sensory-motor connectivity prescribed by a dorsoventral positional template. Cell *147,* 653-665.

Tanabe, Y., William, C., and Jessell, T.M. (1998). Specification of motor neuron identity by the Mnr2 homeodomain protein. Cell *95,* 67-80.

Takamiya, M., and Campos-Ortega, J.A. (2006). Hedgehog signalling controls zebrafish neural keel morphogenesis via its level-dependent effects on neurogenesis. Dev. Dyn. *235,* 978-997.

Waddington, C.H. (1942). The canalization of development and the inheritance of acquired characteristics. Nature *150,* 563-5.

Wendik, B., Maier, E., and Meyer, D. (2004). Zebrafish mnx genes in endocrine and exocrine pancreas formation. Dev. Biol. *268,* 372-383.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. *25,* 1-47.

Xu, Q., Mellitzer, G., Robinson, V., and Wilkinson, D.G. (1999). In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. Nature *399*, 267-271.

Yamada, T., Pfaff, S.L., Edlund, T., and Jessell, T.M. (1993). Control of cell pattern in the neural tube: motor neuron induction by diffusible factors from notochord and floor plate. Cell *73*, 673-686.

Yamaguchi, Y., Shinotsuka, N., Nonomura, K., Takemoto, K., Kuida, K., Yosida, H., and Miura, M. (2011). Live imaging of apoptosis in a novel transgenic mouse highlights its role in neural tube closure. J. Cell Biol. *195*, 1047-1060.

MOVIE CAPTIONS:

Movie S2.1, supporting Figure 2.1B

In toto imaging of neural tube patterning. Red: cell membrane, Blue: cell nuclei, Green: pMNs (*tg(mnx1:gfp)*). Movie is a combination of top view (top) and cross-sectional view (bottom) 3D rendering of the same timelapse dataset. Somites become visible in the top view and notochord is visible soon in the cross-sectional view. Axonal projections of MNs are visible in the end of the cross-sectional movie. The temporal resolution between each frame is 140 seconds. Scale bar: 10µm.

Movie S2.2, supporting Figure 2.1E

Cell movements towards pMN domain formation. Red: cell nuclei from *tg(actb2:h2b-tdTomato)*, Green: pMNs (*tg(mnx1:gfp)*). Movie is a 3D projection of the cross-section view of the original timelapse dataset. Notochord, somite, neural plate and the enveloping layer (EVL) cells can be

distinguished. The GFP expression domain expands dorsally at later stages, and differentiating MNs migrate laterally on both sides. At the end of the movie, the embryo begins wiggling indicating functioning motor circuitry. Scale bar: 10μm.

Movie S2.3, supporting Figure 2.4A

Migration of progenitors and their morphological changes. Red: cell membrane, Gree: pMNs (*tg(mnx1:gfp)*). Movie is a 3D projection of the top view of the neural plate of a mosaic labeled embryo. The neural plate undergoes convergence and extension along with cell intercalation and mixing. The labeled cells show highly dynamic processes in all directions. Under this resolution these projections are rarely found to extend beyond the cell diameter. They also do not appear persistent or correlative to the direction of cell movement. Several cells can be seen turning on transgenic markers and become epithelialized while slowing down. Time resolution is 360 seconds. Scale bar: 10μm.

Movie S2.4, supporting Figure 2.3B,C,C'

Ptch2:Kaede expression patterns at 12hpf and 17hpf. Red: cell membrane, Green: Kaede. Movie is a progressing cross-sectional view (Along the AP axis, P to A, dorsal is up) of the original Z-stacks taken on the neural keel of a 12hpf and a 17hpf *tgBAC(ptch2:kaede)* embryo. At 12hpf, Kaede levels show heterogeneity among the positive cells and do not form a clear ventral to dorsal gradient. Brighter cells can often be found dorsally. At 17hpf, Kaede levels now form a good ventral to dorsal gradient with a triangular negative cell in the middle (the MFP), flanked on both sides by two columns of cells that show strongest Kaede expression (the LFPs). Dimmer cells are found in a graded manner dorsal to the strongest Kaede+ cells.

Movie S2.5, supporting Figure 2.9D

Positional switch of a pMN and a LFP cell. Red: cell membrane, Blue: cell nuclei, Bright green sphere: pMN cell, Orange sphere: LFP progenitor (large) and its daughters (small). Movie is a time sequence of a cross-section plus bird-eye coronal section view (separated by the second green line) rendered by GoFigure2 of the original timelapse dataset. Notochord boundary is visible rising on the left. The pMN cell started closer to the notochord and switched its position with one LFP daughter in the end.

Movie S2.6, supporting Figure 2.10D

Later time window timelapse of *tg(mnx1:gfp)* embryo. Red: cell nuclei, Green: pMNs and MNs (*tg(mnx1:gfp)*). Movie is a 3D projection of the cross-section view of the original timelapse dataset. GFP expression domain shows a dorsal boundary and a ventral boundary, the later bordering 3 floor plate cells, with two LFPs on the side and one MFP in the middle (Figure 2.10C). The locations of these cells are stable in this time window. The temporal resolution between each frame is 240 seconds. Scale bar: 10μm.

Movie S2.7, supporting Figure 2.9F

Misplacement of *mnx1:gfp+* cells under Cdh2 perturbation. Red: dnCdh2-cherry fusion protein, Green: pMNs (*tg(mnx1:gfp)*). Movie is a combination of top view (top) and cross-sectional view (bottom) 3D rendering of the same timelapse dataset. Loss of adhesion prevents cell sorting as the red cells fail to intercalate into the neural keel/tube. The temporal resolution between each frame is 360 seconds. Arrows: Ectopic pMNs. Scale bars: 10μm.

EXTENDED EXPERIMENTAL PROCEDURES:

Zebrafish strains and maintenance

AB wild-type strains were used. *Tg(actb2:h2b-tdTomato), tg(actb2:mem-mCherry)* and *tg(actb2:h2b-mCherry2)* strains were created as described (Kawakami, 2004) using the vector pMTB which contains the minTol2 transposon arms, the β-actin2 enhancer and first intron, an SP6 promoter, a multiple cloning site, and the SV40pA. The transgenes for fluorescent proteins were linked to a sequence coding human histone 2B (H2B) protein. *Tg(GBS:rfp)* line was made by injecting GliBS-mRFPNuc plasmid with I-SceI. Eight copies of Gli binding sites (Sasaki et al., 1997) were cloned in front of the minimal promoter to drive the expression of monomeric RFP (mRFP) tagged with a nuclear localization signal. *Tg(shh:gfp)* (Shkumatava et al, 2004), *tgBAC(ptch2:kaede)* (Huang et al, 2012, NOTE: the original *ptc1:kaede* line was renamed to *ptch2:kaede* to conform with zebrafish nomenclature guidelines of naming zebrafish genes after the human orthologue), *tg(nkx2.2a:mgfp)* (Jessen et al, 1998), *tg(olig2:gfp)* (Shin et al, 2003), *tg(olig2:dsRed)* (Kucenas et al, 2008), *tg(gata2:gfp)* (Ng et al, 2005), *tg(elavl3:kaede)* (Sato et al, 2006, NOTE: the original *huc:kaede* line was renamed to *elavl3:kaede* to conform with zebrafish nomenclature guidelines of naming zebrafish genes after the human orthologue), *tg(mnx1:gfp)* (Flanagan-Steet et al, 2005, obtained from Zebrafish International Resource Center (ZIRC)), *tg(msxb:gfp)* (Keating lab, obtained from ZIRC), and *tg(gfap:gfp)* (Chen et al, 2010, obtained from ZIRC) have been described. Natural spawning was used and time of fertilization was recorded according to the single cell stage of each clutch. Embryos are incubated in 28ºC during imaging and all other times except room temperature during injections and dechorionating. Staging was recorded using morphological criteria and aligned to the normal table (Kimmel et al,

1995). All fish are housed in fully equipped and regularly maintained and inspected aquarium facilities. All related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Microinjections of mRNA and morpholino

pCS or pMTB constructs containing *mem-citrine*, *mem-mCherry*, *h2b-cherry*, *mem-EBFP2* (labeling mRNAs) and zebrafish *mnx1*, *mnx2a* (Wendik et al, 2004, NOTE: the original *mnr2a* gene was renamed to *mnx2a* to conform with zebrafish nomenclature guidelines of naming zebrafish genes after the human orthologue), *dncdh2-cherry* (Rieger et al, 2009) were used for mRNA synthesis with the mMESSAGE mMACHINE system (Ambion). 1-cell stage embryos were injected (Nanoject) 2.3nl 40ng/µl each *h2b-cherry* and *mem-citrine* mRNA and were screened (MVX fluorescent macro scope, Olympus) for health and brightness before imaging. For mosaic labeling, one blastomere of 1 to 32-cell stage embryos was injected with approximately 1nl 20ng/µl labeling mRNA with or without 10ng/µl *mnx1*, *mnx2a* mRNA or 2ng/µl *p53* morpholino (control) or 50µM *cdh2* morpholino (GeneTools, Lele et al, 2002). Variations of concentrations may be used (see corresponding figure legends). One round of immediate screening was applied to mosaically injected embryos to eliminate damaged embryos and/or ones that missed the injection, as identified by lack of retaining Phenol Red (co-injected color label).

Cyclopamine treatment, immunostaining

Cyclopamine (LC laboratories) was dissolved in 100% ethanol to make 50mM stock solution and diluted in embryo water to 100µM immediately prior to treatment. Equal amount ethanol

dilution was prepared for controls. Dechorionated embryos on agarose coated plates were fully submerged in treatment in darkness until mounting for imaging. Nkx6.1 antibody (Developmental studies hybridoma bank, F55A10) and DAPI staining were performed with standard protocols (Available upon request).

Timelapse 2-photon/confocal imaging

Embryos were staged and mounted as described (Kimmel et al, 1995; Megason, 2009), using the dorsal mount with a stereoscope (Leica MZ12.5). Mounting specifics for capturing the neural plate-tube transition are follows: First, dechorionate the embryo prior to mounting on a clean agarose coated plate to avoid damaging the yolk region that is not yet fully covered by epiboly; Second, transfer the embryo with a blunted glass pipette to the mounting dish and use a hair loop to gently rotate the embryo so that its dorsal side is facing up in the middle; Third, gently rotate the embryo in anterior-posterior axis to raise the vegetal axis so that it forms an angle (θ, see Figure 2.2B.v) with the horizontal plane: $\theta=70^{o}$ for 60% epiboly embryos, 30^{o} for 90% epiboly embryos and intermediate angles for stages in between, the angles do not need to be exact but are preferred for best cell tracking coverage; Finally, slowly and gently slide the glass cover slip over the wells while watching the embryo to assure no disturbance. Live imaging was performed using a Zeiss 710 confocal/2-photon microscope and Zen software with a home-made heating chamber maintaining 28ºC and C-Apochromat 40X 1.2 NA objective. Other objectives were also used for non-timelapse imaging. Chameleon laser (Coherent) line 1020nm was used for 3-channel 2-photon in toto imaging and 405nm, 488nm, 514nm, 561nm and 594nm for confocal time courses and other single Z-stacks. Image settings vary by brightness and intended imaging time span of the particular embryo in each experiment. Typically, a bright embryo (labeled with

h2b-cherry and mem-citrine) capable of sustaining trackable signal at 2-minute temporal resolution for >10 hours would be able to provide sufficient signal for clearly distinguishing the deepest cells (such as axial mesoderm cells and yolk syncytial nuclei) at epiboly stages with the imaging parameters listed here: Excitation laser wavelength: 1020nm, output power: 300mW, pixel dwell time: 1μs, line averaging: none. With 40X objectives, the imaging space is: ~0.4μmX0.4μm, and 800X800 pixels per image, with an interval of 1μm through Z for 80 to 150μm, depending on signal strength and regions desired to capture of the future neural tube, in the 80μm setting only the ventral 1/3 of the neural tube by 20hpf can be fully captured. To determine the starting Z location for the embryo so that maximum coverage of the ventral neural tube is achieved, at the starting time point, 40μm of imaging space in Z should be included over the top point of the embryo. In an optimal movie that captures the whole process from 9hpf to 18hpf, starting from earliest detectable Shh responses and ending upon emergence of functional MNs, progenitors of 2-3 full ventral neural segments can be tracked through most time points, yielding an estimated collection of tracks containing 15-20 MFPs, 30-40 LFPs and >60 pMNs. This estimate reflects the potential information content of good *in toto* datasets in this study but is practically difficult to fully extract in analysis (discussed below in Data analysis section). We note that the chance of success in capturing movies of this coverage standard is approximately 30%. The changes accompanying the embryo during this period are dramatic and fast there are many variables that could reduce the chance of success during sample preparation and imaging. For example, the embryos under the scope may behave in several predictable (yet difficult to prevent) ways that cut potential tracking window short (e.g. rotation, titling, etc). We recommend practice of mounting and careful preparations before imaging and using healthy embryos whenever possible to improve success rate. Using faster microscopy set-ups such as light sheet

illumination may improve success rate by allowing a larger imaging space while maintaining trackable temporal resolution.

Data analysis

Raw data from Zen software (Zeiss) are in .lsm format, which were exported to single image files and were pre-processed to GoFigure2 compatible image stacks with custom scripts (available upon request). GoFigure2 (www.gofigure2.org) is an open-source, C++ based, cross-platform software (Microsoft Windows, Mac OS X, Linux), for the analysis of in toto image sets. Tutorials and demos are available at www.youtube.com/user/TheGoFigure2SF. Nuclear segmentation and tracking was performed manually with the GoFigure2 interface using mesh editing and tracking features, fully illustrated here: https://wiki.med.harvard.edu/SysBio/Megason/TrackingIllustration. In tracking segmentation, the center point of a nucleus (e.g. represented by h2b-cherry signal) is clicked once to plant a seed, and GoFigure2 generates a spherical mesh to mark the location of the cell, the meshes of the same cell are connected in the mesh database by a track entry as the segmentation/tracking progresses through time points. Intensity measurement is performed similarly by mesh segmentation inside the cell. To extract the information of the reference point for calculating LM/DV distances, the top center point of the axial mesoderm/notochord is segmented continuously along the anterior-posterior axis, these notochord profile meshes form a curve in the 3D dataset, and the minimal distance of any cell mesh to this curve is computed and defined as the LM/DV distance, these calculations are verified by direct measurement from raw datasets using the 3D distance ruler function in Zen and GoFigure2. Tracks split at cell divisions. The mitotic phase is normally clearly distinguishable over 12 minutes (6 time points). If a mother cell

is seen in metaphase at time point -1, it would be in anaphase at time point 0 with two daughter nuclei but not full membrane separation (cytokinesis), and can thus be unequivocally tracked. The mother and 2 daughter tracks are linked within GoFigure2 to create a lineage. Meshes, tracks, and lineages were exported from GoFigure2 and further processed and plotted with Microsoft Excel. The exported mesh data contains cell coordinates in the 4D dataset, volumes, fluorescent intensities, and cell type annotations. Coordinates were normalized to μm and hpf in space and time for computing distances and speed. For relative speed calculation, first the position vector of the cell P is normalized with the population average position vector Pa by Pr=P-Pa, then relative speed v is given by v=[Pr(t1)-Pr(t0)]/(t1-t0). For population average speed, the absolute values of individual speeds were used. Intensities were normalized by volume; background intensities were measured in cells that are known to be negative for markers examined, for example, notochord cells and somite cells. See also Figure 2.2D, 2.4B-D, and corresponding legends. Other images were extracted from Zen and AxioVision (Zeiss). Movies were generated with GoFigure2, FluoRender2.8 (Wan et al, 2009) and ImageJ (Abramoff et al, 2004).

Cell type identification

In general, cell types are identified using the behavior of transgenic markers throughout the track whenever possible. In high temporal resolution, trackable movies the demand of two bright ubiquitous cell markers (nuclear and membrane) only allows one additional transgenic reporter to be imaged at the same time, while in analysis it is necessary to compare more than one cell type in the same movie. This limitation requires additional cell type identification criteria to be applied. We used information collected from different transgenic reporters and the stability of the

final pattern of the ventral neural tube after 16hpf to register cell types by late locations, in addition, distinct morphologies and axonal projection patterns of some cell types were referred to. Details and related figures are summarized below (other figures use transgenic markers for identifications):

MFPs: medial floor plate cells (*shh:gfp+*) in the early neural tube of fish embryos form a single column of apically constricted cells in the bottom center, as early as 15hpf (Figure 2.3A, Shkumatava et al, 2004; Schafer et al, 2005), sitting directly above the notochord, which has clearly identifiable morphology when notochord cell membranes are labeled. As observed in *tg(shh:gfp)* movies this MFP domain is stable onwards and cells no longer move (Figure 2.10D, data not shown). The apical constriction also makes MFPs triangular shape in cross-sectional view (Compare Figure 2.3A to Figure 2.3B, 2.9C,D). We used both final location and morphology to identify MFPs where *shh:gfp* was not used: Figure 2.3B.iii,iv, Figure 2.9.

LFPs: lateral floor plate cells (*nkx2.2a:mgfp+*) in the early neural tube of fish embryos form 2 columns of apically constricted cells flanking the MFPs, together they form the floor plate of trunk neural tube, 3 cells ventral to the *mnx1:gfp+* cells (Figure 2.10A, Schafer et al, 2005, Huang et al, 2012). Most LFPs settle by 16hpf and remain stable onwards (Figure 2.6C, 2.10B-D), therefore their negative *mnx1:gfp* expression and ventral localization to the *mnx1:gfp+* domain provide reliable identification. This identification was used in Figure 2.9. The LFPs are also the strongest *ptch2:kaede* expressing cells, together with their locations beside the MFP (Figure 2.2A, Movie S2.4), they can be identified as in Figure 2.3B.iii,iv.

KA"s: Kolmer-Agduhr neurons (*gata2:gfp+* in the LFP domain, Huang et al, 2012) are born within the LFP domain. We tracked these cells as a control for LFP domain stability and reliability of LFP identification by location after 16hpf (Figure 2.10B-D). We further confirmed

these cells are actual neurons by their projections and co-expression of *elavl3:mKate* (data not shown).

pMNs: motorneuron progenitor cells (*olig2:gfp*+ and *mnx1:gfp*+) form a stripe-like domain in the early neural tube of fish embryos. Note that weak *mnx1:gfp* expression extend dorsally beyond pMN domain in the final pattern, and weak *olig2:gfp* expression can be seen in the LFPs (Figure 2.2A). Therefore we used *mnx1:gfp* expression to identify the pMNs dorsal to the floor plate cells, while *olig2:gfp* expression to identify the pMNs ventral to the p2s. In tracks reported in Figure 2.9, which used *mnx1:gfp* to identify pMNs, most pMNs were further followed to 19hpf when they begin to differentiate to MNs and extend axonal projections to verify correct identification. The second strongest *ptch2:kaede* expression and immediate dorsal localization to the LFPs in the final pattern (Figure 2.2A, Movies.S2.4) also allows pMN identification as in Figure 2.3B.iii,iv.

MNs: motorneurons (*mnx1:gfp*+, strong) differentiate from the pMN domain starting around 17-18hpf in the imaged zone and show characteristic cell body rounding up and lateral migration. We used axonal projections at the ventral root of the spinal cord of these cells for identification (Westerfield et al, 1986; Figure S2.3B, data not shown). We also further confirmed these cells are actual neurons by their projections and co-expression of *elavl3:mKate* (data not shown).

p2s and V2s: ventral 2 progenitor and interneurons (*gata2:gfp*+, labeling V2 interneurons, Figure 2.2A) were tracked in several *gata2:gfp* movies and contributed to lineage tracing data (Figure 2.7A, 2.10B-D). These tracks were not used to compare domain formation.

CHAPTER 2 ADDITIONAL REFERENCES:

Abramoff, M.D., Magalhaes, P.J., Ram, S.J., (2004). Image Processing with ImageJ.

Biophotonics Int. *11(7),* 36.

Charron, F., Stein, E., Jeong, J., McMahon, A.P., and Tessier-Lavigne, M. (2003). The morphogen sonic hedgehog is an axonal chemoattractant that collaborates with netrin-1 in midline axon guidance. Cell *113,* 11-23.

Chen, W., Burgess, S., and Hopkins, N. (2001). Analysis of the zebrafish smoothened mutant reveals conserved and divergent functions of hedgehog activity. Development *128,* 2385-2396.

Chen, H.L., Yuh, C.H., and Wu, K.K. (2010) Nestin Is Essential for Zebrafish Brain and Eye Development through Control of Progenitor Cell Apoptosis. PLoS One *5(2)*: e9318.

Kawakami, K. (2004). Transgenesis and gene trap methods in zebrafish by using the Tol2 transposable element. Methods Cell Biol. *77,* 201-222.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. Dev. Dyn. *203,* 253-310.

Kimura, Y., Satou, C., and Higashijima, S. (2008). V2a and V2b neurons are generated by the final divisions of pair-producing progenitors in the zebrafish spinal cord. Development *135,* 3001-3005.

Kucenas, S., Takada, N., Park, H.C., Woodruff, E., Broadie, K., and Appel, B. (2008) CNS-derived glia ensheath peripheral nerves and mediate motor root development. Nat. Neurosci. *11(2)*:143-151.

Lewis, K.E., Eisen, J.S. (2001) Hedgehog signaling is required for primary motoneuron induction in zebrafish. Development *128(18)*:3485-95.

Megason, S.G. (2009). In toto imaging of embryogenesis with confocal time-lapse microscopy. Methods Mol. Biol. *546*, 317-332.

Ng, A.N., de Jong-Curtain, T.A., Mawdsley, D.J., White, S.J., Shin, J., Appel, B., Dong, P.D.,

Stainier, D.Y., and Heath, J.K. (2005). Formation of the digestive system in zebrafish: III. Intestinal epithelium morphogenesis. Dev Biol *286,* 114-135.

Park, H.C., Mehta, A., Richardson, J.S., and Appel, B. (2002). Olig2 is Required for Zebrafish Primary Motor Neuron and Oligodendrocyte Development. Dev. Biol. *248,* 356-368.

Sasaki, H., Hui, C., Nakafuku, M., and Kondoh, H. (1997). A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro. Development *124,* 1313-1322.

Sato, T., Takahoko, M., and Okamoto, H. (2006). HuC:Kaede, a useful tool to label neural morphologies in networks in vivo. Genesis *44,* 136-142.

Wan, Y., Otsuna, H., Chien, C.B., and Hansen, C. (2009). An interactive visualization tool for multi-channel confocal microscopy data in neurobiology research. IEEE Trans. Vis. Comput. Graph. *15,* 1489-1496.

Westerfield, M., McMurray, J.V., and Eisen, J.S. (1986). Identified motoneurons and their innervation of axial muscles in the zebrafish. J. Neurosci. *6,* 2267-2277.

Chapter 3

Genetic Regulatory Noise in the Interpretation of Shh Signaling by Neural Progenitors

Fengzhu Xiong, Andrea R. Tentner, Sean G. Megason

Department of Systems Biology, Harvard Medical School, Boston, 02115, USA

Author Contributions: F.X., A.R.T. and S.G.M. conceived this study. F.X. and A.R.T. performed the experiments and analyzed the data. F.X. wrote this manuscript under the guidance of S.G.M..

SUMMARY

Developmental patterning relies on interpretation of inductive signals (morphogens) by intracellular gene regulatory networks (GRNs). It has been proposed that the GRN underlies the accuracy and robustness of observed target gene expression patterns. Recent evidence suggests cellular mechanisms directly ensure patterning precision in dynamic tissues, raising the possibility that highly precise GRNs are unnecessary in such systems. Here we image the GRN dynamics in single neural progenitors during Shh signaling in zebrafish embryos. We found the average Shh response dynamics are distinct between progenitor groups of different final fates. However, at the single cell level there is a high degree of variability in Shh response and GRN dynamics within a fate class and overlap in response between classes. Further heterogeneity is found in the timing of response initiation. Our results reveal a basic limit of accuracy in the GRN of Shh interpretation. Together, we discuss the spatial, temporal, and GRN sources of noise and suggest a general constraint on the precision of signal interpretation in the patterning of dynamic

tissues.

INTRODUCTION

Cells receive signals from their surroundings and respond with gene expression changes. Developing tissues use this strategy to inform cells of their positions and thus organize the formation of patterns. The reproducibility and precision of many patterns are seemingly in contrast with the noisy molecular reactions and dynamic cellular and tissue movements accompanying morphogenesis. Several models have been proposed to explain the observed accuracy of signal interpretation and patterning. First, cells may integrate signals temporally or spatially to average out fluctuations (Dessaud et al., 2007; Zhang et al., 2012; Di Talia and Wieschaus, 2012; Little et al., 2013). Second, the intracellular gene regulatory network (GRN) may provide interpretation logic that translates noisy, continuous signal exposure into discrete, stable downstream gene expression states (Balaskas et al., 2012). Third, cellular level mechanisms, such as cell sorting and chemotaxis, provide a means to correct pattern imprecision after signal interpretation has taken place (Kay and Thompson, 2009; Xiong et al., 2013).

These mechanisms may act together or one may be dominant, depending on the properties of the system. Notably, in dynamic tissues such as the zebrafish early neural tube, where cells undergo a lot of movement and proliferation, the spatial patterns of signaling and fate specification are initially noisy but are then corrected by cell sorting (Xiong et al., 2013). This raises the question of where the noise comes from in this system. It is possible that the GRN provides a precise interpretation for each cell but cell movements or imprecision of Shh gradient introduce the observed spatial noise (extracellular origin of noise). It is also possible that a fundamental limit of GRN accuracy exists and the GRN contributes to noisy specification

(intracellular origin of noise). Such noise-limited accuracy in signal transduction is now well appreciated in other systems but has been difficult to address in vertebrate development (Paulsson 2004; Eldar and Elowitz, 2010; Balazsi et al., 2011).

Here we explore the GRN input-output function in single neural progenitors *in vivo* by measuring their Shh signaling dynamics and downstream reporter expression at the same time, using triple color imaging combined with mosaic cell labeling. Using cell tracking, we found that the populations of different progenitors respond to Shh with distinct dynamics, consistent with previous tissue level studies in many systems. Surprisingly, a significant degree of unpredictability exists at the single cell level: cells that respond to Shh similarly may produce distinct downstream reporter dynamics and adopt different fates, whereas cells that share fate reporter dynamics may have distinct Shh dynamics. Furthermore, cells show fate-independent heterogeneity of timing in initiating Shh response that spans 3 hours. Together, our data show that in addition to noise caused by cell mixing, neural progenitors also show significant temporal and GRN noises in Shh interpretation. Based on these observations we suggest a general constraint in morphogen patterning that drives evolution towards different mechanisms for precision depending on the degree of tissue dynamics.

RESULTS

Analyzing signaling response and downstream dynamics in single cells *in vivo*

We have previously established a powerful *in toto* imaging platform to capture the full neural tube patterning process in zebrafish embryos at cellular resolution (Xiong et al., 2013), in which we used both nuclear and membrane fluorescent proteins of different colors (cell markers) to label and follow all cells. However, only one single additional transgenic reporter can be used

with this strategy. This is due to the requirement that the two cell markers need to be very bright for the high temporal sampling rate required for cell tracking, which causes bleed-through signal to interfere with the detection of usually much weaker transgenic reporters (especially when they just start to be expressed). In our hands, with bleed-through from two bright cell markers there is only enough spectral range for a single transgenic reporter. In addition, high temporal resolution can photobleach the signal of weak transgenic reporters. To overcome these problems, here we image two transgenic reporters at once and employ only one cell tracking label (e.g. mem-EBFP2) and a lower temporal resolution. This strategy improves the sensitivity to detect reporters and reduces reporter bleaching. However, to ensure that cells can still be tracked accurately given the lower temporal resolution, only a fraction of cells were labeled by mosaic injections (Figure 3.1A). This approach results in a lower number of trackable cells per dataset as compared to *in toto* movies, but allows us to track information flow through the GRN in single cells *in vivo*. We chose *tg(ptch2:kaede)* as the Shh response reporter and combined it with other downstream GFP reporters. Using this method, we imaged mem-EBFP2 as a separate track to both capture cell outlines and convert Kaede protein to red fluorescence. Acquiring the Kaede(red) and GFP signals in another track prevented the interference of widespread autofluorescence introduced by the 405nm laser used to excite EBFP2. These experiments led to 15 trackable datasets of different downstream reporters (Figure 3.1A, Movie S3.1). We followed the identified cells using GoFigure2 to quantify cell position and fluorescence levels at all time points as described previously (Xiong et al., 2013). These tracks allow a direct quantitative comparison of reporters on a per-cell basis (Figure 3.1B), offering an opportunity to understand the relationship between GRN components in pairs *in vivo* (Figure 3.1C).

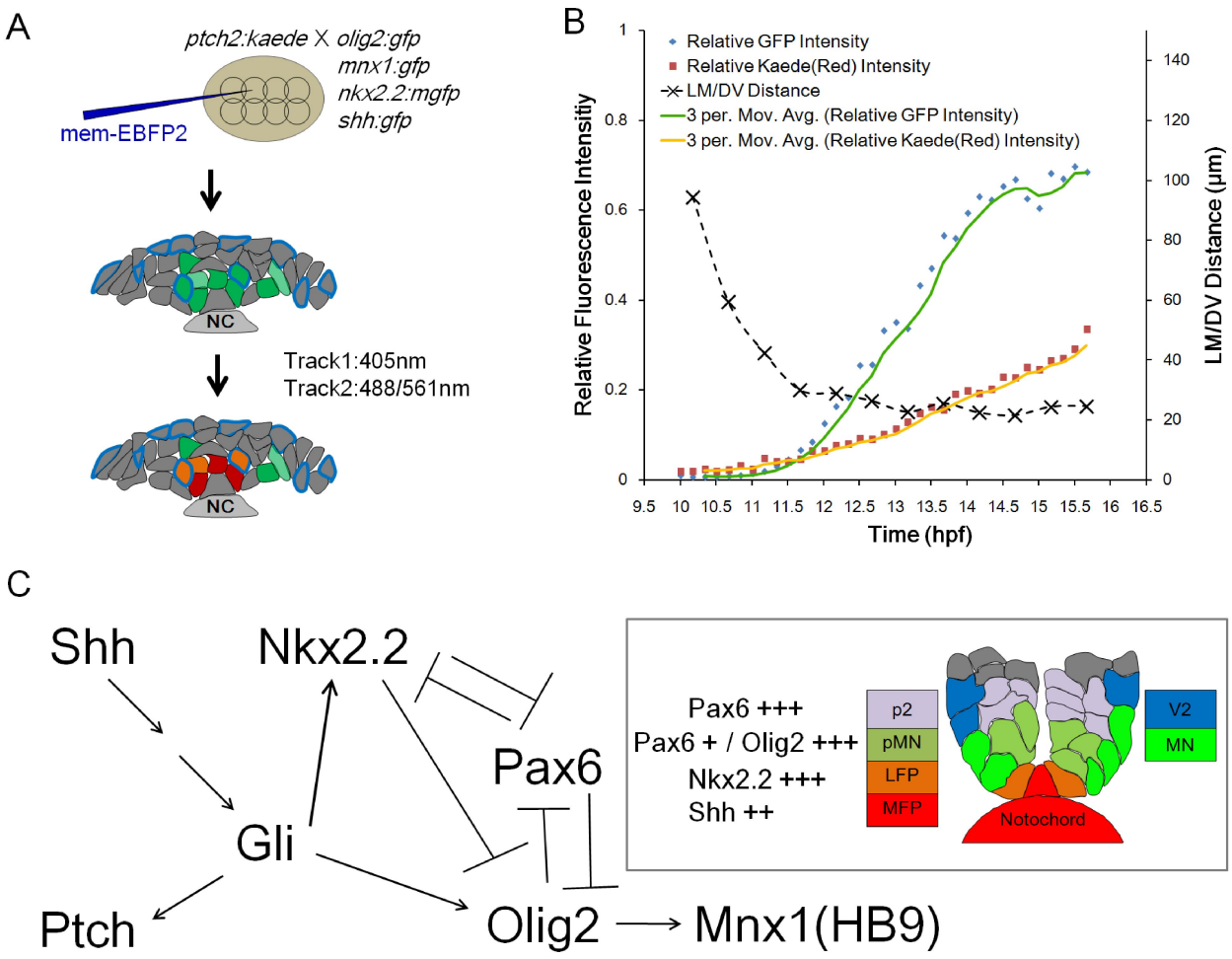


Figure 3.1. Strategy of triple color mosaic imaging and cell tracking for neural progenitor GRN dynamics.

(A) Schematic illustration of double reporter imaging and tracking method. Embryos from transgenic crosses of *ptch2:kaede* and a fate reporter line is labeled mosaically with mem-EBFP2, which is imaged while Kaede is converted to red in the first track for each Z-stack. See also Movie S3.1.

(B) Example track containing the expression dynamics of both reporters and the cell's movement trajectory represented by its lateral-medial/dorsal-ventral distance to the center top of the notochord (Xiong et al., 2013). The intensity values were adjusted by removing bleed-through and normalization using saturation intensities in raw images (data now shown).

Figure 3.1 (*Continued*) (C) Schematic illustration of the Shh interpretation GRN that leads to different ventral neural progenitor fate. Arrows indicate activation and T-shapes indicate repression. Inset box shows the correlation between gene expression and fate and the Dorsal-Ventral locations of different progenitors. p2, V2 interneuron progenitor. pMN, motroneuron progenitors. LFP, lateral floor plate cells (homologous to p3 domain in chick and mice). MFP, medial floor plate cells. See also Balaskas et al., 2012; Xiong et al.,2013.

Trajectories of reporter expressions in different neural progenitor types

We collected cell tracks from each dataset and compared the dynamic patterns of reporters with final positions and fates. *Tg(mnx1:gfp)* shows strong expression when Shh response is low or moderate (Figure 3.2A), these cells are found to be motorneuron progenitors (pMNs) or more dorsal progenitors to the pMNs. Cells that are further dorsal show no response to Shh or expression of *mnx1:gfp*. Cells that are ventral to the pMNs may respond strongly (the lateral floor plates, LFPs), or weakly (the medial floor plates, MFPs). Notochord cells, as expected, do not respond to Shh signaling. These results are largely consistent with an accurate GRN model (only extracellular-origin noises), although some cells show very similar Kaede profiles but distinct GFP dynamics (e.g. Compare the 3rd pMN and the 4th LFP in Figure 3.2A).

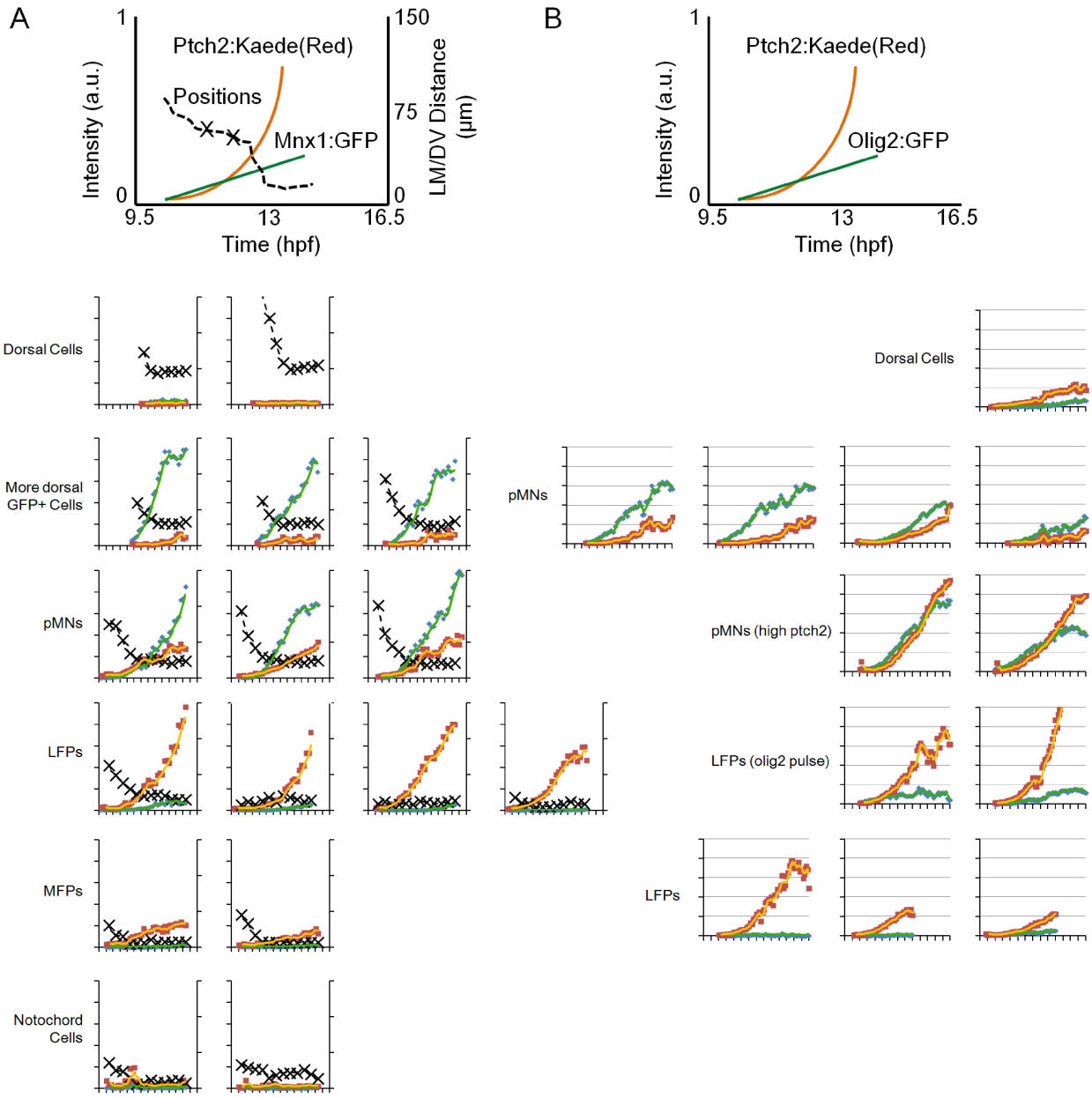


Figure 3.2. Single cell reporter trajectories.

(A) Trajectories from *ptch2:kaede* X *mnx1:gfp*. The top diagram is an example track plot that provides annotations of the axes and trajectories. Individual cell tracks were plotted below and are organized vertically from top to bottom by cell types (Annotated on the right for each row). The annotations were removed due to space limitations. pMNs: motorneuron progenitors. LFPs: lateral floor plate cells. MFPs: medial floor plate cells. Same for (B).

Figure 3.2 (*Continued*) (B) Trajectories from *ptch2:kaede* X *olig2:gfp*.

A key player in the Shh interpretation GRN is Olig2 (Figure 3.1C, Balaskas et al., 2012). Compared to Mnx1, Olig2 is a more upstream transcription factor that biases cells towards pMN fates. Olig2 is repressed by Nkx2.2 (a LFP marker) at higher Shh concentrations and exposure times, which causes a gene expression change from olig2 to nkx2.2 (Dessaud et al., 2007). We analyzed a *tg(olig2:gfp); tg(ptch2:kaede)* dataset to assess the relationship between Olig2 and Ptch2 dynamics (Figure 3.2B). As expected, more dorsal cells that have low response do not express GFP, and the cells that have moderate Shh response express GFP and become pMNs. However, the distinction in Shh response dynamics between the GFP- and GFP+ cells is not apparent (Compare Dorsal cell with the 2nd pMN in Figure 3.2B). Surprisingly, we also found GFP+ cells that have high Shh response (3rd row in Figure 3.2B). The response level is comparable to those of the more ventral LFPs. In addition, we indeed found evidence of gene expression change. Some LFPs (4th row in Figure 3.2B) show onset of GFP which then plateaus, indicating a transient pulse of Olig2 transcription activity. Interestingly, this pattern is not universal in LFPs, some LFPs appear directly specified without a pulse of GFP (5th row in Figure 3.2B). These data show that again, Shh response levels largely correlate with Olig2 behavior, however, individual cells show a high degree of variability and the same fates may be achieved through distinct GRN dynamics.

Noise in the correlation between fate and signaling response in single cells

To quantitatively compare the dynamics of Shh response in the tracked cells, we looked at the measured reporter dynamics in each track, whose rate of increase in fluorescence levels

indirectly reflects the transcription rate of the gene. We applied a linear fit to the *ptch2:kaede* reporter expression in its first 4 hours after response initiation. The ratio of saturated voxels in this time window is low and the increase is close to linear (data not shown), therefore the linear fit provides a good measure of the rate of increase (the slope) and initiation time of increase (the time axis intercept). We then listed the slope values of the fit and marked identified cells correspondingly (Figure 3.3A). Analysis of 70 tracks show that different cell fates cluster at distinct slope values (Figure 3.3B), but a lot of overlap exists near the boundaries, suggesting cells receiving similar Shh signaling at those levels may choose either fate (Figure 3.3A).

We have shown some cells directly become LFPs without turning on a transient pulse of olig2 expression. To further test the gene expression change model, we asked when the cells start responding to Shh, by looking at the linear fit intercept on the time axis of our tracks (Figure 3.3C), which is indicative of the time the cell starts to respond to Shh. If LFPs are specified by more extended Shh exposure than pMNs, one may expect to see LFPs start responding to Shh earlier than pMNs. However, we found early and late responding cells of both high and low slopes, spanning a 3 hour window. We note that the anterior-posterior difference (where more anterior cells on average start responding earlier than posterior cells) does not fully account for this heterogeneity, as the tracked cells are located within a narrow range of approximately 4 somites.

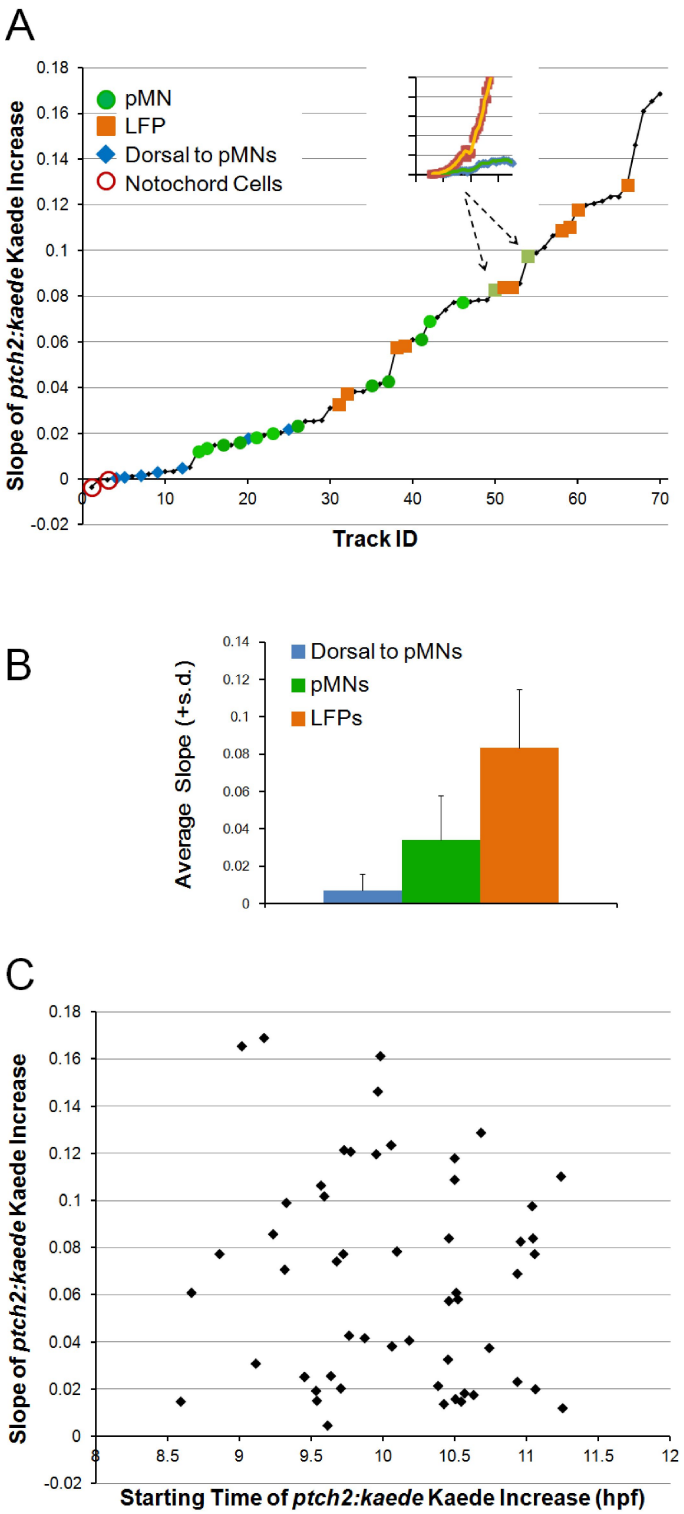


Figure 3.3. Temporal and GRN dynamics heterogeneity among neural progenitors

(A) Responding levels to Shh signaling in single cells. A linear fit to the *ptch2:kaede* reporter

Figure 3.3 (*Continued*) expression trajectory in its first 4 hours after its initiation is applied to the tracks as shown in Figure 3.2. The slope value of the fitted line represents how fast the Kaede level increases, indicative of how strongly the cell is responding to Shh. 70 tracks from 4 datasets were plotted sequentially by the slope value. Cells of identified fates were highlighted by corresponding markers. Insert and arrows show two LFPs cells that transiently expressed *olig2:gfp* as found in Figure 3.2B.

(B) Average ± s.d. of slope values by cell fates. The differences are significant under a t-test: p(blue-green)=0.01; p(green-orange)=3e-4. n=8 for dorsal cells, n=14 for pMNs and n=12 for LFPs.

(C) Temporal heterogeneity in Shh response initiation. Each marker represents the time (hpf) where the fitted line from (A) intercepts the time axis. For cells that did not initiate the response (e.g. notochord cells) this time value is significantly out of range and is not plotted.

Together, these data show that neural progenitors respond to Shh signaling at different times and levels, and that final fate choices are correlated with the signal level on average but not the onset time. Importantly, single cells may show different fate choices after very similar Shh responses, suggesting that the GRN interpretation of Shh morphogen is neither precise nor stereotypic.

DISCUSSION

Our findings have several interesting implications. First, the presence of GRN noise raises the question of how progenitor fates are eventually determined. One possibility is that regulation further downstream of the Olig2-Nkx2.2 GRN consolidates fates. Consistent with this hypothesis,

the precision of correlation in Mnx1:gfp tracks appears better (compare Figures 3.2A and 3.2B) than the Olig2:gfp tracks. More cell tracks are required for a quantitative comparison. Other possibilities include inherent heterogeneity in naive progenitors or inputs from other pathways during Shh interpretation. We note that it is also possible that these specification events are naturally not deterministic, and a binary fate choice was made in part by chance. Second, we found that the correlation between GRN behavior and fate is not deterministic, raising the question of whether the different routes matter if they apparently lead to the same fate. One possibility is that different GRN behaviors correlate to subtype identities in the pools of the same fate. More tracks will be needed to measure the proportions of each GRN signature in the population and further tracking with additional sub-fate markers may reveal the correlations. Alternatively, the GRN signature might not matter (e.g., a pMN resulting from high Shh response undergoes further development indistinguishable from a pMN that went through low Shh response). Finally, the presence of a cell sorting mechanism allows the noise to be canalized and pattern precision ensured (Figure 3.4A). We note that the position used in our studies are nucleus/cell centers but the overall cell morphology and cellular processes such as cytonemes should be considered (Sanders et al., 2013). The signal measurement we used is from reporters. While BAC reporters are reasonably accurate, they are still one step away from the actual functional fusion proteins from the endogenous locus (Trinh et al., 2011).

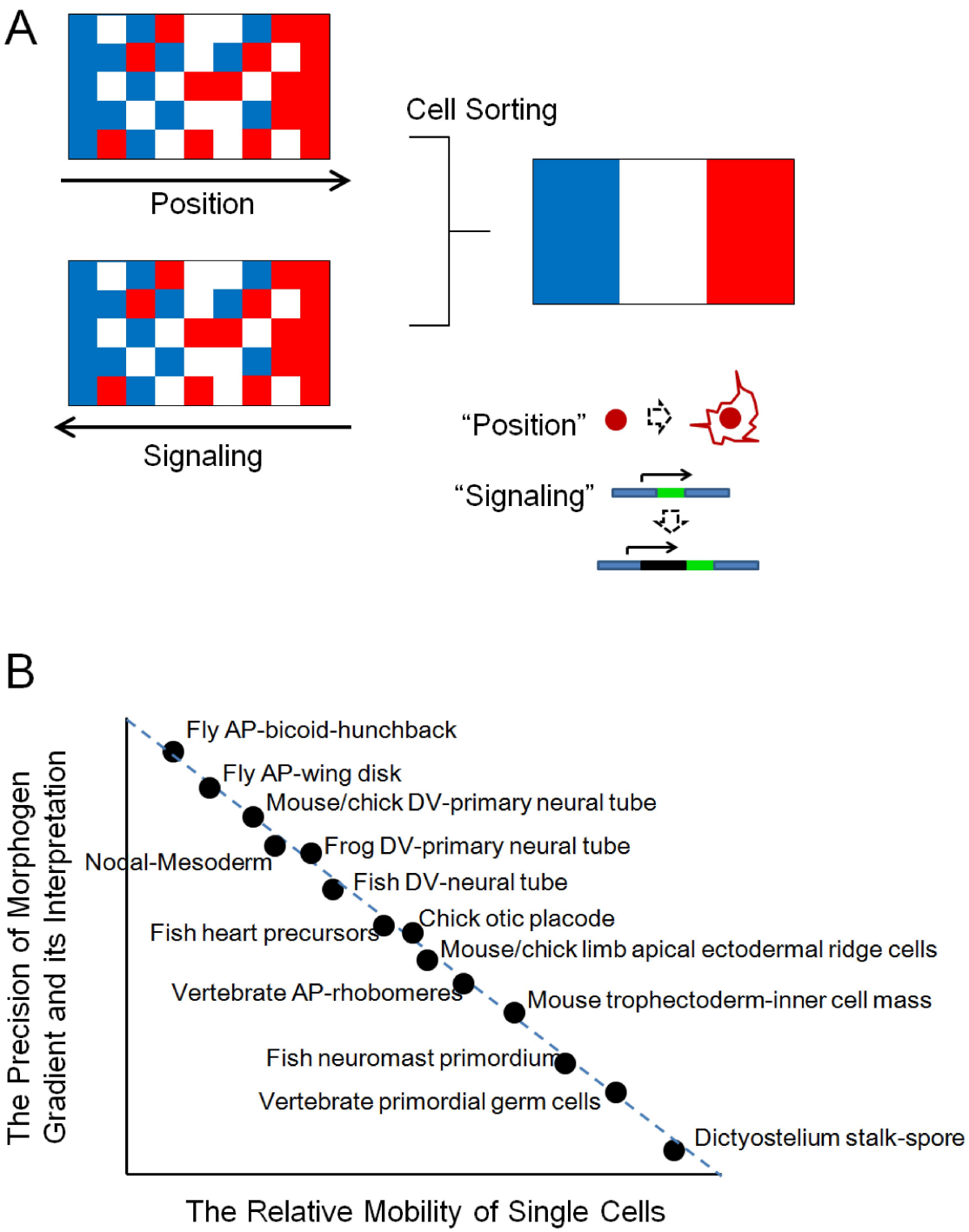


Figure 3.4. A possible general constraint by cell mobility and GRN precision

(A) Both positional information and its interpretation are noisy, requiring cell sorting mechanism to ensure patterning precision.

(B) A possible general constraint of cell movement on the precision of morphogen gradient and

Figure 3.4 (*Continued*) its interpretation. The precision of morphogen can be defined as the percentage of positional overlap of cell fates across the spatial domain boundaries. The relative mobility of cells can be defined the cell's displacement in relation to the population average over the time scale of specification. The dots were planted on the plot approximately drawing on following references (ordered corresponding to the dots from top left to bottom right): Gregor et al., 2007; Hamaratoglu et al., 2013; Balaskas et al., 2012; Schier 2005; Edlund et al., 2013; Xiong et al., 2013; Keegan et al., 2005; Streit 2002; Altabef et al., 1997; Guo et al., 2003; Xu et al., 1999; Plusa et al., 2008; Matsuda et al., 2013; Dormann et al., 2003; Kay and Thompson, 2009).

Our thorough analysis of spatial (Xiong et al., 2013), temporal and GRN noise (this study) in Shh mediated ventral neural tube pattern formation allows us to speculate on mechanisms for precise patterning on a broader scale. The aforementioned mechanisms, either the GRN or cell sorting, are all arguably costly to the embryo. On one extreme (e.g. Bicoid morphogen gradient interpretation in fly embryos) where there is little cell movement, the GRN must evolve to a precision that can distinguish very small differences in signal gradients (Gregor et al., 2007). While it is unclear how that precision is achieved, it is possible that the mechanism is delicate and not very evolvable due to its high sensitivity and dependence on a precise gradient. On the other extreme (e.g. Dictyostelium slug formation) where no inductive gradient is needed and cells completely rely on lateral communication and sorting to assemble a precise pattern (Kay and Thompson, 2009), it is conceivable that cells will use a large number of surface molecules to communicate and energy to move, which is expensive. These examples suggest that, when one mechanism is in place and effective (e.g. cell sorting in the fish neural tube), others may be

allowed to be sloppy for the overall efficiency. This means that evolution may drive morphogen interpretation to be more precise when movement is limited. In the opposite direction, when morphogen interpretation is physically impossible to be precise (e.g. dynamic tissue), cellular mechanism must arise (Figure 3.4B). As long as pattern precision is functionally important and evolutionarily constrained, we propose that evolution will follow the Pareto front in Figure 3.4B to produce a range of GRN and cellular mechanisms that underlie the apparent diversity for a precise pattern to be achieved.

EXPERIMENTAL PROCEDURES

Zebrafish strains and maintenance

See Xiong et al., 2013 for protocols, sources and references for transgenic strains used in this study. All fish related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Microinjections of mRNAs

For mosaic injections, one blastomere of 8 to 32-cell stage embryos was injected with approximately 1nl 20ng/µl one labeling mRNA (h2b-cherry, mem-citrine, mem-cherry, mem-EBFP2 and combinations).

Timelapse 2-photon/confocal imaging

Live imaging was performed using a Zeiss 710 confocal/2-photon microscope (objective: C-Apochromat 40X 1.2 NA) with a home-made heating chamber maintaining 28ºC. See Xiong et al., 2013 for details.

Image Data analysis

Cell segmentation and tracking was performed using GoFigure2, an open-source, cross-platform

software application we have developed for image analysis (www.gofigure2.org). Segmentation and track tables exported from GoFigure2 were further processed and plotted with Microsoft Excel. Linear fit of tracks were performed using the linear fitting function in Matlab (Mathworks). See also Xiong et al., 2013.

ACKNOWLEDGEMENTS

The authors thank D. D'India for fish care, K. Mosaliganti for data analysis, H. Otsuna for movie making, T. Mitchison, A. Schier, A. Kicheva, J. Briscoe and Megason lab members for discussions. This work is supported by NIH grants HG004071 and DC010791. F. X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

CHAPTER 3 REFERENCES

Altabef, M., Clarke, J.D., and Tickle, C. (1997). Dorso-ventral ectodermal compartments and origin of apical ectodermal ridge in developing chick limb. Development *124,* 4547-4556.

Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K.M., Briscoe, J., and Ribes, V. (2012). Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell *148,* 273-284.

Balazsi, G., van Oudenaarden, A., and Collins, J.J. (2011). Cellular decision making and biological noise: from microbes to mammals. Cell *144,* 910-925.

Dessaud, E., Yang, L.L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitch, B.G., and Briscoe, J. (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature *450,* 717-720.

Di Talia, S., and Wieschaus, E.F. (2012). Short-term integration of Cdc25 dynamics controls mitotic entry during Drosophila gastrulation. Dev. Cell. *22,* 763-774.

Dormann, D., and Weijer, C.J. (2003). Chemotactic cell movement during development. Curr. Opin. Genet. Dev. *13,* 358-364.

Edlund, A.F., Davidson, L.A., and Keller, R.E. (2013). Cell segregation, mixing, and tissue pattern in the spinal cord of the xenopus laevis neurula. Dev. Dyn.

Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. Nature *467,* 167-173.

Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. (2007). Probing the limits to positional information. Cell *130,* 153-164.

Guo, Q., Loomis, C., and Joyner, A.L. (2003). Fate map of mouse ventral limb ectoderm and the apical ectodermal ridge. Dev. Biol. *264,* 166-178.

Hamaratoglu, F., de Lachapelle, A.M., Pyrowolakis, G., Bergmann, S., and Affolter, M. (2011). Dpp signaling activity requires Pentagone to scale with tissue size in the growing Drosophila wing imaginal disc. PLoS Biol. *9,* e1001182.

Kay, R.R., and Thompson, C.R. (2009). Forming patterns in development without morphogen gradients: scattered differentiation and sorting out. Cold Spring Harb Perspect. Biol. *1,* a001503.

Keegan, B.R., Feldman, J.L., Begemann, G., Ingham, P.W., and Yelon, D. (2005). Retinoic acid signaling restricts the cardiac progenitor pool. Science *307,* 247-249.

Little,S.C., Tikhonov, M., Gregor, T. (2013). Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. Cell *154,* 789-800.

Matsuda, M., Nogare, D.D., Somers, K., Martin, K., Wang, C., and Chitnis, A.B. (2013). Lef1 regulates Dusp6 to influence neuromast formation and spacing in the zebrafish posterior lateral line primordium. Development *140,* 2387-2397.

Paulsson, J. (2004). Summing up the noise in gene networks. Nature *427,* 415-418.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135,* 3081-3091.

Schier, A.F. (2009). Nodal morphogens. Cold Spring Harb Perspect. Biol. *1,* a003459.

Streit, A. (2002). Extensive cell movements accompany formation of the otic placode. Dev. Biol. *249,* 237-254.

Sanders, T.A., Llagostera, E., and Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. Nature *497,* 628-632.

Trinh le, A., Hochgreb, T., Graham, M., Wu, D., Ruf-Zamojski, F., Jayasena, C.S., Saxena, A., Hawk, R., Gonzalez-Serricchio, A., Dixson, A., *et al.* (2011). A versatile gene trap to visualize and interrogate the function of the vertebrate proteome. Genes Dev. *25,* 2306-2320.

Xiong, F., Tentner, A.R., Huang, P., Gelas, A., Mosaliganti, K.R., Souhait, L., Rannou, N., Swinburne, I.A., Obholzer, N.D., Cowgill, P.D., Schier, A.F., and Megason, S.G. (2013). Specified neural progenitors sort to form sharp domains after noisy Shh signaling. Cell *153,* 550-561.

Xu, Q., Mellitzer, G., Robinson, V., and Wilkinson, D.G. (1999). In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. Nature *399,* 267-271.

Zhang, L., Radtke, K., Zheng, L., Cai, A.Q., Schilling, T.F., and Nie, Q. (2012). Noise drives sharpening of gene expression boundaries in the zebrafish hindbrain. Mol. Syst. Biol. *8*, 613.

MOVIE CAPTIONS:

Movie S3.1. Supporting Figure 3.1.

Movie S3.howing example dataset. This embryo is from transgenic crosses of *ptch2:kaede* (Shown as blue signal in movie) and *mnx1:gfp* (Shown as green signal in movie) injected with mem-EBFP2 (Shown as red membrane signal in movie). Movie is a combination of top view (top) and cross-sectional view (bottom) slices of the same timelapse dataset, the slices follow a single tracked neural progenitor (White dot). Time stamp is in hh:mm format. Temporal resolution is 6min/stack. Scale bar: 20µm.

Chapter 4

Interplay between oriented cell divisions and geometrically/physically constrained cell shapes underlies morphogenesis of simple epithelia

Fengzhu Xiong, Wenzhe Ma, Kishore R. Mosaliganti, Andrea R. Tentner, Nicolas Rannou, Arnaud Gelas, Lydie Souhait, Ian A. Swinburne, Nikolaus D. Obholzer, Marc W. Kirschner, Sean G. Megason

Department of Systems Biology, Harvard Medical School, Boston, 02115, USA

Author Contributions: F.X. and S.G.M. conceived this study. F.X. performed most experiments and analyzed the data. F.X. and W.M. built the model. K.R.M., N.R., A.G., L.S. contributed to data analysis. A.R.T performed the cell dissociation experiments. I.A.S and N.D.O provided reagents. F.X. and S.G.M. wrote the manuscript. M.W.K and S.G.M supervised this study.

SUMMARY

The development and function of simple epithelia rely on the formation of specific cell shapes (squamous, cuboidal, columnar). This morphogenesis process is regulated by geometry, mechanics and cell behaviors that extensively interact. However, an integrated model to address this complexity has been illusive due to the scarcity of quantitative data. Here we measure and

model surface morphogenesis of early zebrafish embryos. We show that dynamics of surface area, cell volume and cell number quantitatively correlate with the flattening of surface cells. Using cell tracking, we found cell shape determines division orientation and thus feeds-back to regulate surface cell number. Using modeling we found that cell mechanical interactions define observed shape distributions. By comparison with other systems and perturbations, we found that division orientation control and mechanical regulation can explain the diversity and dynamics of cell shapes. Together, we propose a general model of epithelial morphogenesis controlled by the interplay of local cell behaviors under global constrains.

HIGHLIGHTS

- Quantitative description of presumptive enveloping-layer (EVL) morphogenesis in zebrafish embryos
- Surface cell number is influenced by cell shape distribution through division orientations
- A mechanically regulated interplay model recapitulates morphogenesis and division dynamics
- Varying division orientation and adhesion can explain diversity of epithelial morphologies and perturbed dynamics

INTRODUCTION

How cells generate forms is a fundamental question in developmental biology. During the patterning and morphogenesis of simple epithelia, cells on the surface of tissues differentiate to

carry out distinct functions including protection, support, selective permeability and secretion. Simple epithelia are categorized by the shape of their cells - e.g. squamous, cuboidal and columnar - which are also usually indicative of their functions. In addition, these cell shapes serve as the basis for more elaborate tissue structures to form (Salazar-Ciudad et al., 2009; Eiraku et al., 2011; Wang et al., 2012), and failure to correctly produce or maintain them causes many defects and diseases such as neural tube defects (NTDs) and various cancers (Deramaudt and Rustgi, 2005; Ciruna et al., 2006; Thiery et al., 2009). A range of geometrical, mechanical, molecular and cell behavioral factors have been recognized to influence the morphogenesis of tissue surfaces towards different epithelial types. First, the expansion of the epithelial tissue enlarges the exposed (free) surface area of cells making them more squamous, whereas constriction of tissue surface reduces cell surface area to make cells more columnar (Kane et al., 2005; Martin et al., 2010; Sato and Clevers, 2013). Second, the growth and proliferation of surface cells make them more columnar whereas volume depletion by cell extrusion, apoptosis, internalization or asymmetric divisions promotes more squamous morphologies (Wolpert and Gustafson, 1961; Chalmers et al., 2003; Tamulonis et al., 2011; Eisenhoffer et al., 2012). Third, the mechanical forces between and within cells (e.g. adhesion and cortical tension), cell polarity cues and extracellular matrix molecules locally regulate cell shapes (Lecuit and Lenne, 2007; Kafer et al., 2007; Fernandes et al., 2010; Guillot and Lecuit, 2013).

Despite previous work on various aspects of surface morphogenesis, it is still unclear how the aforementioned properties of the surface cell layer interact as a whole system to generate different epithelial morphologies. As an integrated process the changes of many factors take

place concurrently. Their effects on the surface cells are intertwined with mechanical feedbacks (Nelson et al., 2005; Ciruna et al., 2006; Gibson et al., 2011). However, the machineries underlying these complex interactions remain poorly understood and it is challenging to investigate all players at the same time. For example, how does a proliferating surface balance the increasing number of cells to maintain an overall squamous morphology? How do mechanical interactions between cells that affect cell shapes locally influence the tissue morphology? What are the key parameters whose regulation could change the direction or dynamics of morphogenesis (e.g. from columnar to squamous)? How may these interactions increase robustness or evolvability of the system?

Quantitative models based on high resolution imaging data (Keller 2013) may provide a means to address these questions at a cell and tissue level, while coarse-graining out the unresolved molecular mechanisms (Megason et al., 2011). Recent advances in imaging allow cell-tracking-based simultaneous measurement of many parameters at enhanced spatial-temporal coverage and resolution, thus providing the required data for model development and validation (Keller et al., 2008; MacMahon et al., 2008; Oliver et al., 2011; Fernandez et al., 2010; Osterfield et al., 2013; Xiong et al., 2013). Theoretical models can include many contributing factors at many different scales thus allowing for a more systematic analysis. Despite inaccuracies that result from assumptions and simplifications, models are able to describe and inform the main tissue constraints and cell behaviors, and further suggest the mechanisms by which morphogenesis progresses (Odell et al., 1981; Gibson et al., 2011; Krupinski et al., 2011; Tamulonis et al., 2011). Importantly, non-intuitive predictions may be drawn and used to guide experimental tests, and

parameter analysis may reveal key differences that produce the apparent variety.

Here we use *in toto* imaging (Megason and Fraser, 2003) to capture the dynamics of surface morphogenesis in early zebrafish embryos. We simultaneously and systematically measured the dynamics of tissue surface area, cell shape, cell volume, cell number, and division orientation. We integrated them to establish a quantitative correlation with the morphology of simple epithelia. By cell tracking we found that surface cell number changes through differentially oriented cell divisions. The division orientation is in turn predicted by the cell shape. In addition, modeled cell mechanics produce the cell shape distribution. Combining these findings we developed an interplay model which recapitulates observed morphogenesis. Further parameter analysis of the model suggests that division orientation control and mechanical regulation can produce the diversity and dynamics of epithelial morphogenesis, which we validated by comparing with other systems and perturbations. Together, our work provides a quantitative description of surface morphogenesis in zebrafish embryos and an insightful analytical platform for its dynamics. We propose that morphogenesis of simple epithelia is achieved by the interplay of local cell behavior and global constraints.

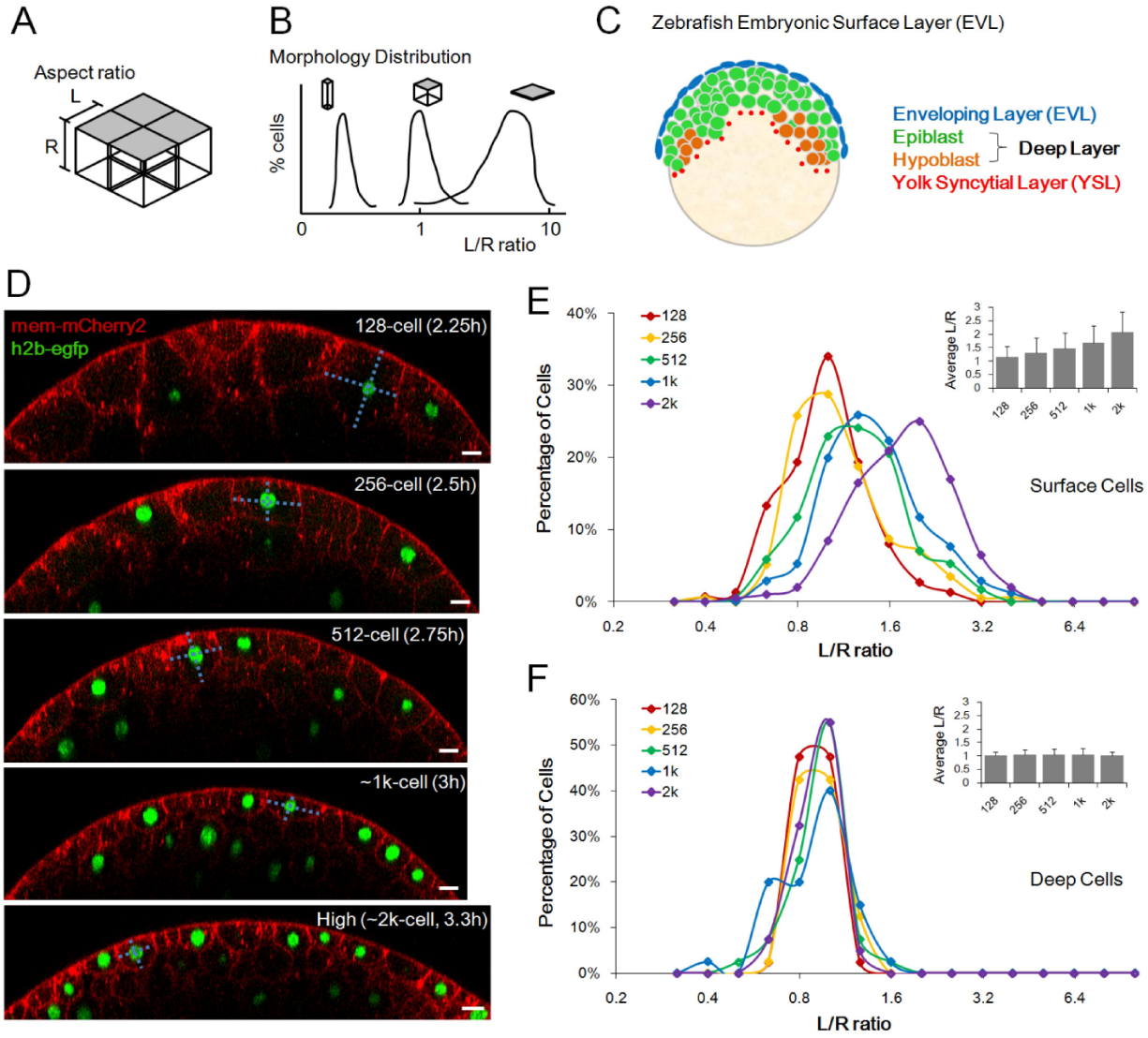


Figure 4.1. Quantitative description of surface layer morphogenesis of zebrafish embryos

(A) Schematic illustration of simplified surface monolayer, grey indicates the free surface, cell shape is represented by aspect ratio L/R.

(B) The morphology of surface layers is represented as a distribution of L/R ratios.

(C) Zebrafish surface layer (the enveloping layer, EVL, blue) system in the context of the whole embryo. Sketch represents the lateral cross-section of a shield stage embryo.

(D) Timelapse imaging data and measuring surface cell L/R ratios across the nucleus (Cyan

Figure 4.1 (*Continued*) lines). Scale bars: 20μm. See also Figures 4.2A-D.

(E) Morphogenesis of surface cells over 5 cell cycles. n=860. Inset: Average L/R±s.d.

(F) Deep cells remain isometric at the same time. n=200.

RESULTS

A general framework for describing surface layer morphogenesis and zebrafish imaging data

Despite the variety of simple epithelia, the types of cell shape changes during their morphogenesis are limited. To simplify measurements and comparisons, we represented cell shapes by a simple parameter: the ratio of length scales of the cell's lateral (along the surface) and radial (perpendicular to the surface) dimensions (L/R, Figure 4.1A). The dynamics of the population can thus be described as a temporal evolution of a distribution of L/R ratios of a certain number of cells (Figure 4.1B). These simplifications allow an intuitive, quantitative representation of surface morphogenesis, capturing not all but an essential component of the shape changes of the cells.

Using direct live imaging of early zebrafish embryos labeled with nuclear and membrane fluorescent proteins (Figures 4.2A-C), we followed the presumptive enveloping layer (EVL, Figure 4.1C), a surface layer that develops into a simple squamous epithelium (Kimmel et al., 1995). The surface cells divide meta-synchronously between 128-cell and high stage over 5 cell cycles (Movie S4.1, Keller et al., 2008; Oliver et al., 2011), producing a course of step-wise morphological changes for the whole layer. We measured morphologies of cells at the time point centered between 2 consecutive cell divisions when the embryonic surface is the smoothest and

most stable (Figures 4.1D,4.2D). The measured shape distributions of the surface cells show a flattening shift (to the right) with time and a wide range of cell shapes at all time points (Figure 4.1E), whereas the deep cells under the surface stay isometric and comparably tightly distributed (Figure 4.1F). Intriguingly, this flattening process happens earlier than known lineage restriction or marker expressions of the EVL (Figures 4.2E-F, Ho 1992; Sagerström et al., 2005). These results are consistent with previous observations of this system (Kimmel and Law, 1985; Kane et al., 1992; Zhang et al., 2011). Thus the evolution-of-distribution representation we use validly describes the process, allowing us to ask how surface morphogenesis happens in a quantitative manner.

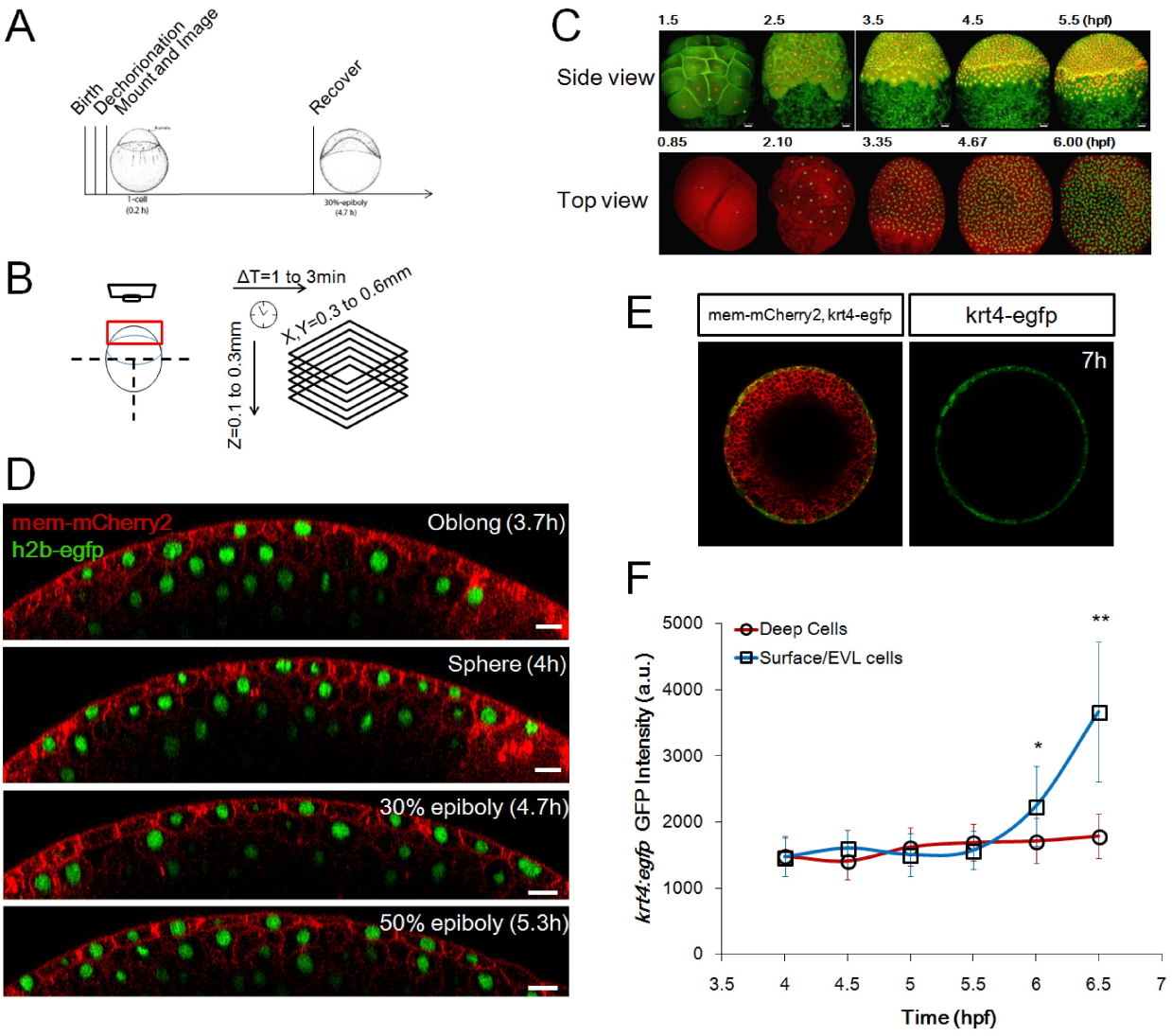


Figure 4.2. Surface imaging of zebrafish embryos and EVL differentiation

(A) Schematic illustration of imaging timeline and protocol. Embryo illustrations courtesy of Kimmel et al., 1995.

(B) Imaging set-up and spatial-temporal coverage and resolution. See also Extended Experimental Procedures.

(C) Example time courses focusing on a lateral view and a top view, respectively. The lateral view images are reused from a movie published with Xu et al.,2012.

Figure 4.2 (*Continued*) (D) Extended timelapse imaging data following Figure 4.1D. Scale bars: 20μm.

(E) *tg(krt4:egfp:caax)* EGFP expression in the differentiated EVL cells (Krens et al., 2011).

(F) Time course of *tg(krt4:egfp:caax)* EGFP intensity in surface cells and deep cells. n=~70 cells per time point measured. At 5.5h p=0.08; *p=5e-6; **p=1e-14 (2 tailed t-test).

Tissue surface area, cell volume and cell number constrain the average surface cell shape

To understand how cell shape may become more flattened or columnar over time, we assessed the dynamics of key parameters that are associated with the morphology of the surface cells, namely the total surface area (**A**) of the layer, surface cell volume ($\mathbf{V_C}$) and surface cell number ($\mathbf{N_S}$). Intuitively, these parameters relate to the L/R ratio of the cells in the following ways: first, the average surface area of cells satisfies: $A/N_S=L^2$; second, the volume of a cell is approximately $V_C=L^2R$. These allow the L/R ratio to be expressed in terms of A,V_C and N_S. While the actual average L/R ratio may be different depending on the type of distribution of cell shapes (SI text.1), the ratio $L/R=(A/N_S)^{1.5}/V_C$ (eq.1) serves as a close estimate of the actual average L/R.

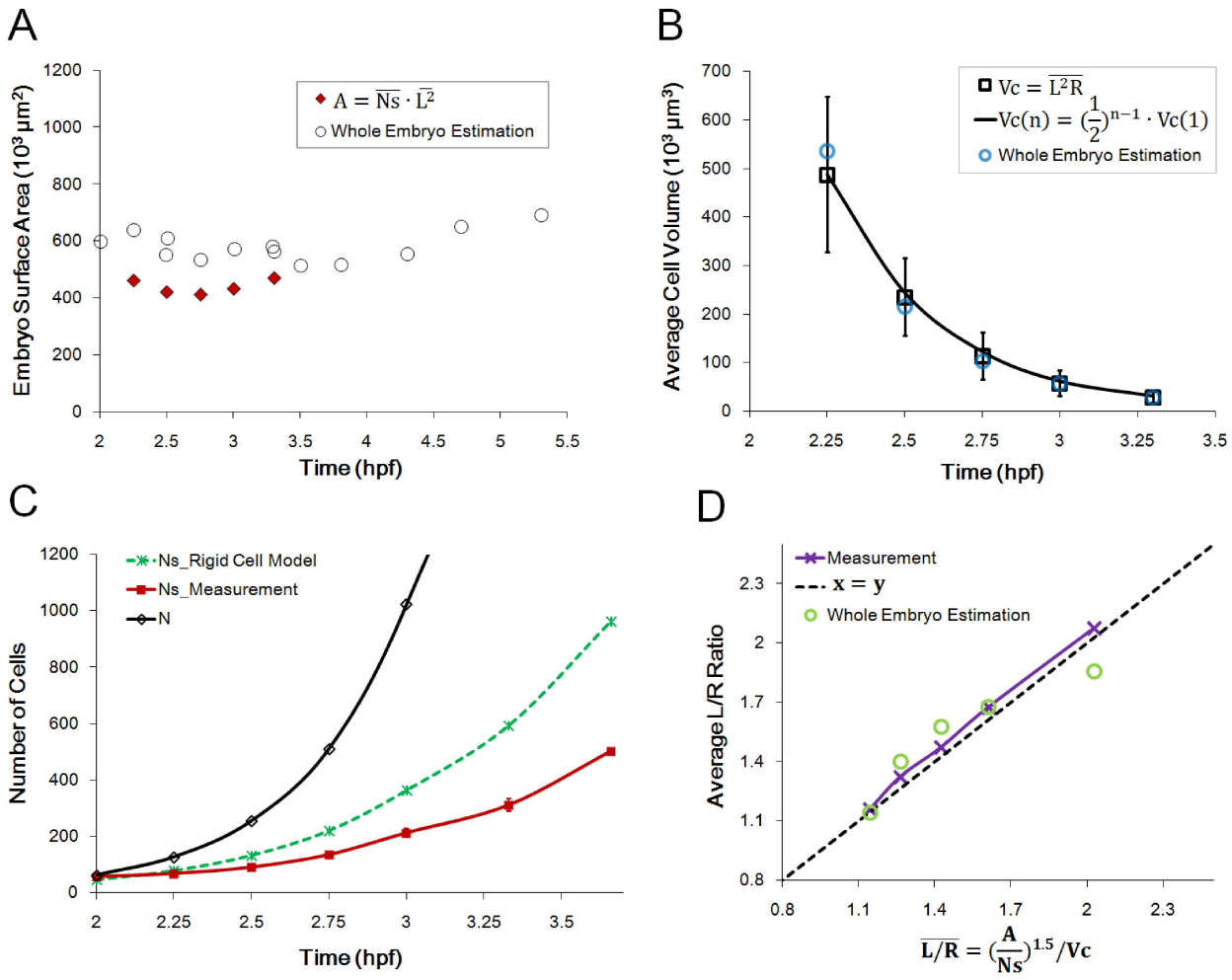


Figure 4.3. Surface area, cell volume and cell number dynamics constrain morphogenesis

(A) Surface area (A) changes of the whole layer. Both estimation methods show changes less than 10% in the 5 cell cycles imaged. The difference between the 2 methods is likely related to the inverse trapezoid shape of the cells so that L is smaller than the surface width. See also Figures 4.4A-B, Extended Experimental Procedures.

(B) Average cell volume (V_C) changes over time. Error bars: s.d. of L^2R estimation. See also Figures 4.4A, C-E.

(C) Surface cell number dynamics. See also Figures 4.4F-G, SI text.2.

(D) Calculated average L/R ratio using eq.1 compared to measurement (from Figure 4.1E) and

Figure 4.3 (*Continued*) whole embryo estimation compiling numbers from (A),(B) and (C). See also SI text.1.

With eq.1, we first measured A both macroscopically (Figure 4.4A) and using calculated average L^2 from single cell measurements (Figure 4.4B). The results indicate that A is stable with under 10% fluctuations between the 5 time points of L/R measurements (Figure 4.3A). Thus, unlike later epiboly stage (Kane et al., 2005), the expansion of surface area is not the cause of cell flattening in this time window. Next, we measured V_C using global average (Figure 4.4A), average L^2R (Figure 4.4C) and full membrane segmentations (Figure 4.4D, Movie S4.2). These different methods agree with each other well and show that the average V_C closely follows an ideal exponential decrease curve (Figure 4.3B) by becoming half at every cell cycle, indicating no cell growth during this time. Individual cells half their volumes roughly equally at divisions (7.3%±5.2% difference from average V_C of two daughters, n=43 pairs; Figure 4.4E, Olivier et al., 2010). According to eq.1, a halving V_C would double L/R after each cell cycle if A and N_S do not change. Since A is stable, the observed slow increase in average L/R (Figure 4.1E) can only be accounted for by an increase of N_S. To test this, we measured N_S after each cell cycle in 8 embryos (Figure 4.4F, Keller et al., 2010), and found that N_S shows a highly consistent increasing trend (Figure 4.4G). Interestingly, this trend is slower than predicted by a model (SI text.2) where surface cells are rigid and do not change morphology (Figure 4.3C), suggesting that flattening *must* happen. Indeed, by combining A, V_C and N_S measurements we obtain an increasing trend of L/R that is consistent with whole embryo estimation, and is similar to, but

does not exactly match the actual average L/R values (Figure 4.1D), indicating the distribution of cell shapes is not symmetric around the average.

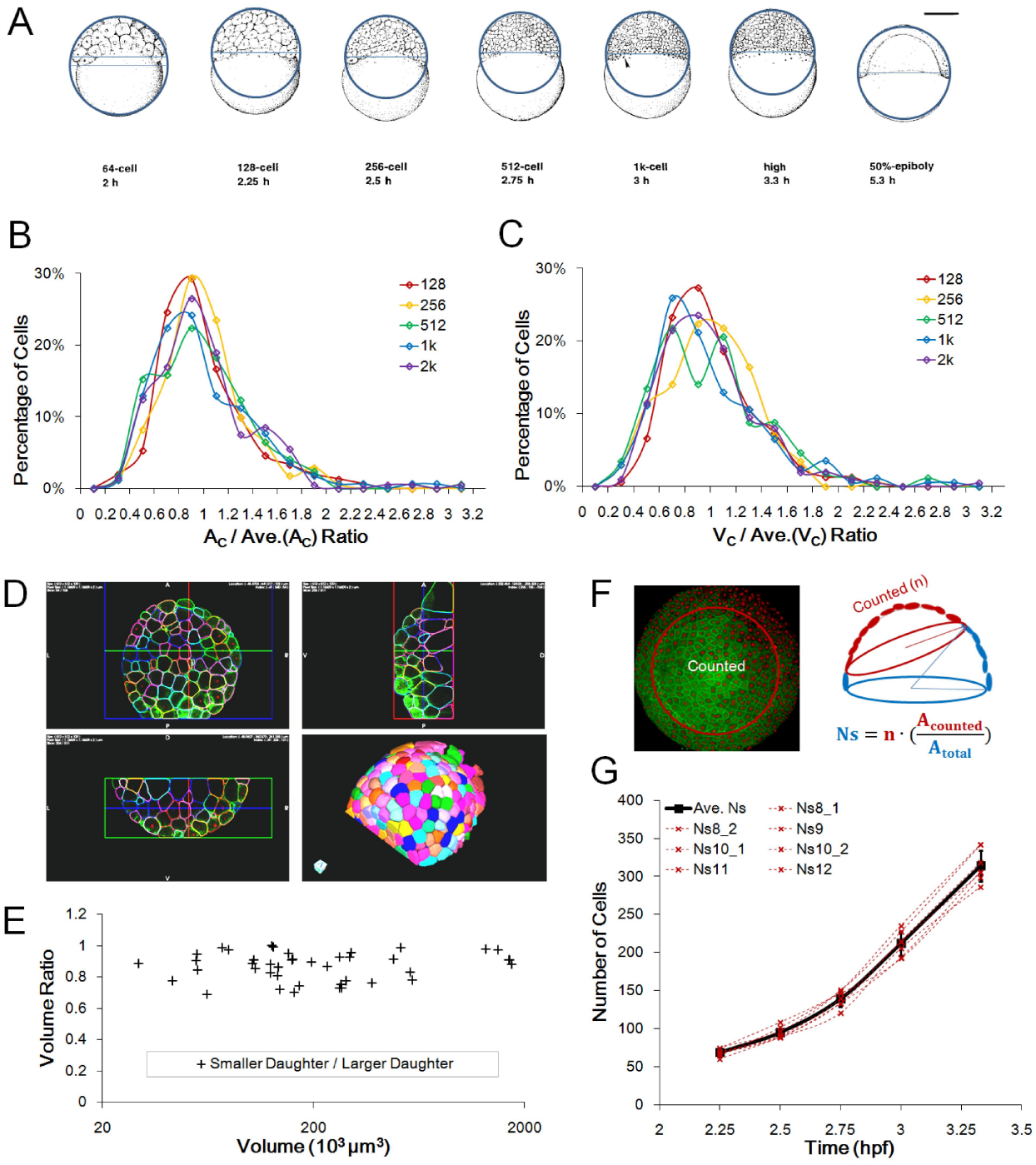


Figure 4.4. Measurements and distributions of surface area, cell volume and number

(A) Schematic illustration of macroscopic surface area and volume measurement using whole

Figure 4.4 (*Continued*) embryo lateral images. Both illustrations and actual photographs are used. The image is fit with a circle to approach the surface shape and a string is drawn on the vegetal edge of surface cells. The surface area and volume of the resulting spherical cap provides rough whole embryo estimate values of A and V_C. Embryo illustrations courtesy of Kimmel et al., 1995.

(B) Normalized distribution of L^2.

(C) Normalized distribution of L^2R.

(D) Examples of full membrane segmentation to measure V_C using ACME and GoFigure2. See also Movie S4.2.

(E) Ratio of sister cell volumes (larger/smaller). Data points used full membrane segmentation measurements as in (D).

(F) Schematic illustration of surface cell counting method. Due to the spherical curvature and imaging limitations only partial surfaces were acquired thus a N_S estimation was used by drawing a circle on 3D perspective views of the surface. The total number equals the cell number in the circle multiplies the ratio between surface area of the circled cap and whole surface area.

(G) Individual N_S counts of 8 embryos and average. The average values were used in modeling. N_S 12 is obtained from a movie by Keller et al., 2010.

Together, these results show that the formation of simple epithelial tissue of a given morphology obeys the geometrical constraints of surface area, cell volume and cell number changes. In this system in particular, the N_S parameter changes non-intuitively and explains the positive shift of the average L/R ratio of surface cells. How, then, is N_S regulated in this system?

Cell shape is predictive of division orientation and its distribution controls N_S

To understand how N_S is regulated to lead to a simple epithelium of a specific cell morphology, we tracked individual cells through time to follow their divisions and the resulting locations of their daughters. Two distinct types of divisions are found to occur in the surface population that affect N_S: the surface-surface (S-S) division that produces two surface daughters (Figure 4.5A), and surface-deep (S-D) division that produces one retaining surface daughter and one deep daughter (Figure 4.5B). S-S division roughly halves the L/R ratios of daughters (Figure 4.5C) whereas S-D division roughly doubles the surface daughter's L/R ratio, while the deep daughter becomes isometric (Figure 4.5D). These results show that S-S divisions contribute to crowding while S-D divisions contribute to flattening. Thus the fraction of these division types at each cell cycle contributes to N_S changes and may influence the average L/R.

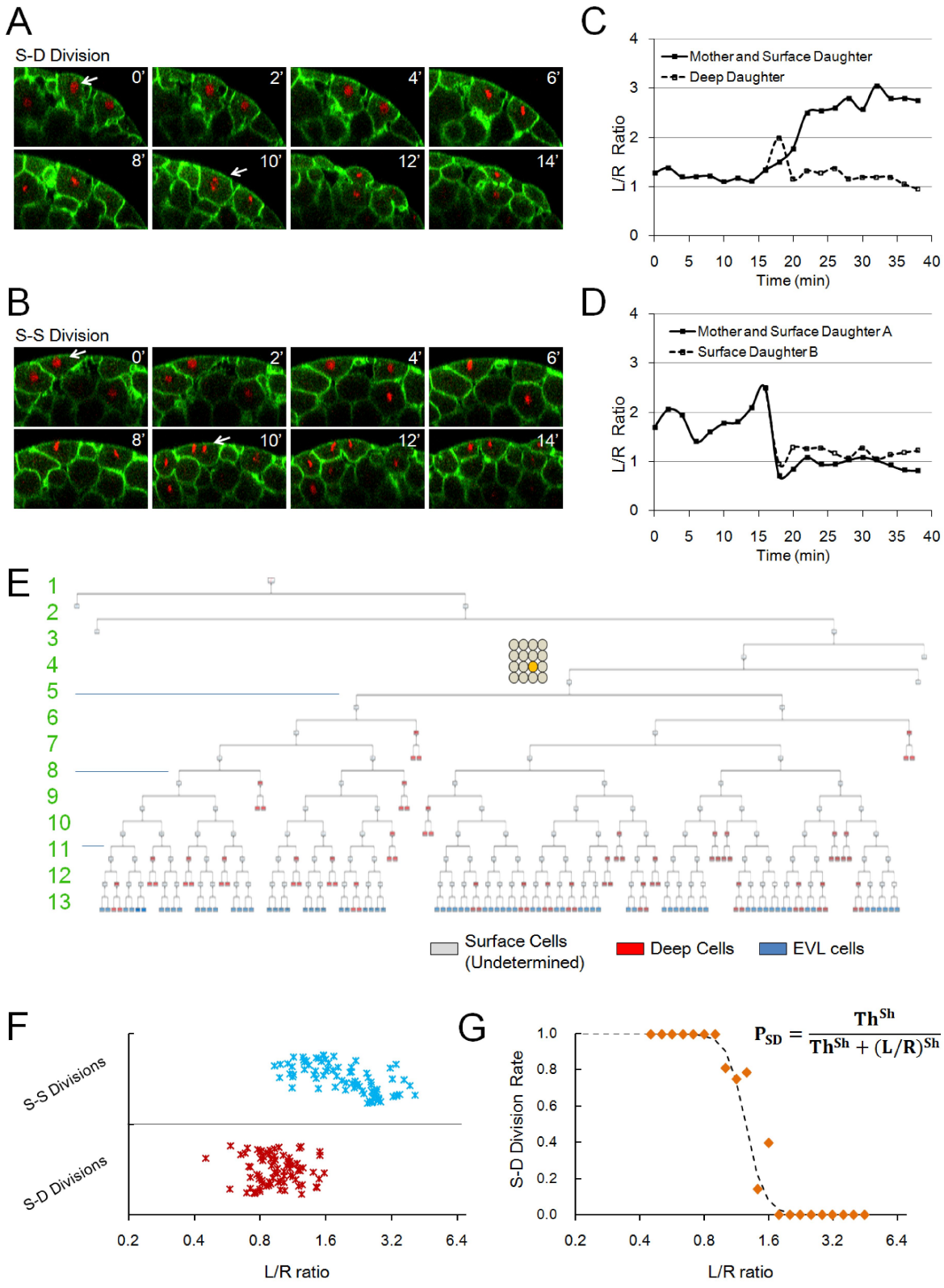


Figure 4.5. Division patterns determine N_S and cell morphology is predictive of division

Figure 4.5 (*Continued*) **orientation**

(A,B) Time courses of example S-D, S-S divisions, respectively. Arrows indicate the tracked cells.

(C,D) Example S-D, S-S mother and daughter cell shape dynamics throughout the cell cycle, respectively.

(E) Full lineage up to 13th cell division of a middle blastomere at 16-cell stage. For simplicity, deep cell branches are not drawn as the divisions always make more deep cells. Numbers indicate cell cycle number.

(F) Relationship between cell shape and division choice. Each marker represents a tracked cell. n=162. See also Figures 4.6A-B.

(G) S-D division rate calculated by partitioning data in (F) according to L/R values. A Hill function switch is used to approach data points. Th=1.25, Sh=10.

To assess other possibilities of N_S changes and understand how a choice of S-S vs. S-D division is made, we first reconstructed full lineage trees of surface cells by tracking future EVL cells from cell cycle 1 to 13 (Figure 4.5E; data not shown). The tree reveals a few interesting properties of N_S dynamics. First, we found no instances of cell extrusion from the surface to out of the embryo or internalization from the surface to deep without division. Second, deep cells only make more deep cells and do not return to the surface. Thus, the change of N_S only occurs through cell division. Third, S-D divisions start to appear at the 6th cell cycle and stop at the 13th when the EVL lineage becomes separate from the deep cells at early epiboly (Ho 1992). Fourth,

S-D divisions may appear at different branch levels and their appearance does not follow a fixed lineage pattern. These results show that although S-S and S-D divisions change N_S and predictably affect L/R of daughters, the history of divisions in a lineage is variable and not predictive of the next division type. Other factors besides lineage history must be considered to understand how division orientation is determined.

An intriguing possibility is the shape of the cells themselves. It has long been recognized that cell shape correlates with division orientation (Hertwig 1884) and several molecular mechanisms have been proposed to account for this relationship (Grill and Hyman, 2005; Kunda and Baum, 2009; Wuhr et al., 2010; Minc et al., 2011). To test this in our system, we tracked 162 cells at different times to their divisions after measuring their L/R ratio. Indeed, highly flattened cells always undergo S-S divisions and highly columnar cells always undergo S-D divisions, while at intermediate aspect ratios a zone of overlapping choice exists (Figure 4.5F). We calculated the fraction of S-D divisions as a function of L/R and fitted it with a Hill function switch (Figure 4.5G). The transition to favor S-S division occurs when the L/R increases over a Threshold (**Th)** of 1.25, and is remarkably sharp (Sharpness (**Sh**) is 10). Consistent with cell tracking, the centrioles align along the future division axes (Figure 4.6A). Furthermore, when cell shape is forcefully changed by injection of an oil droplet, the division orientation aligns with the new long axis (Figure 4.6B). These data establish a quantitative division rule for the surface cells using L/R ratio as input.

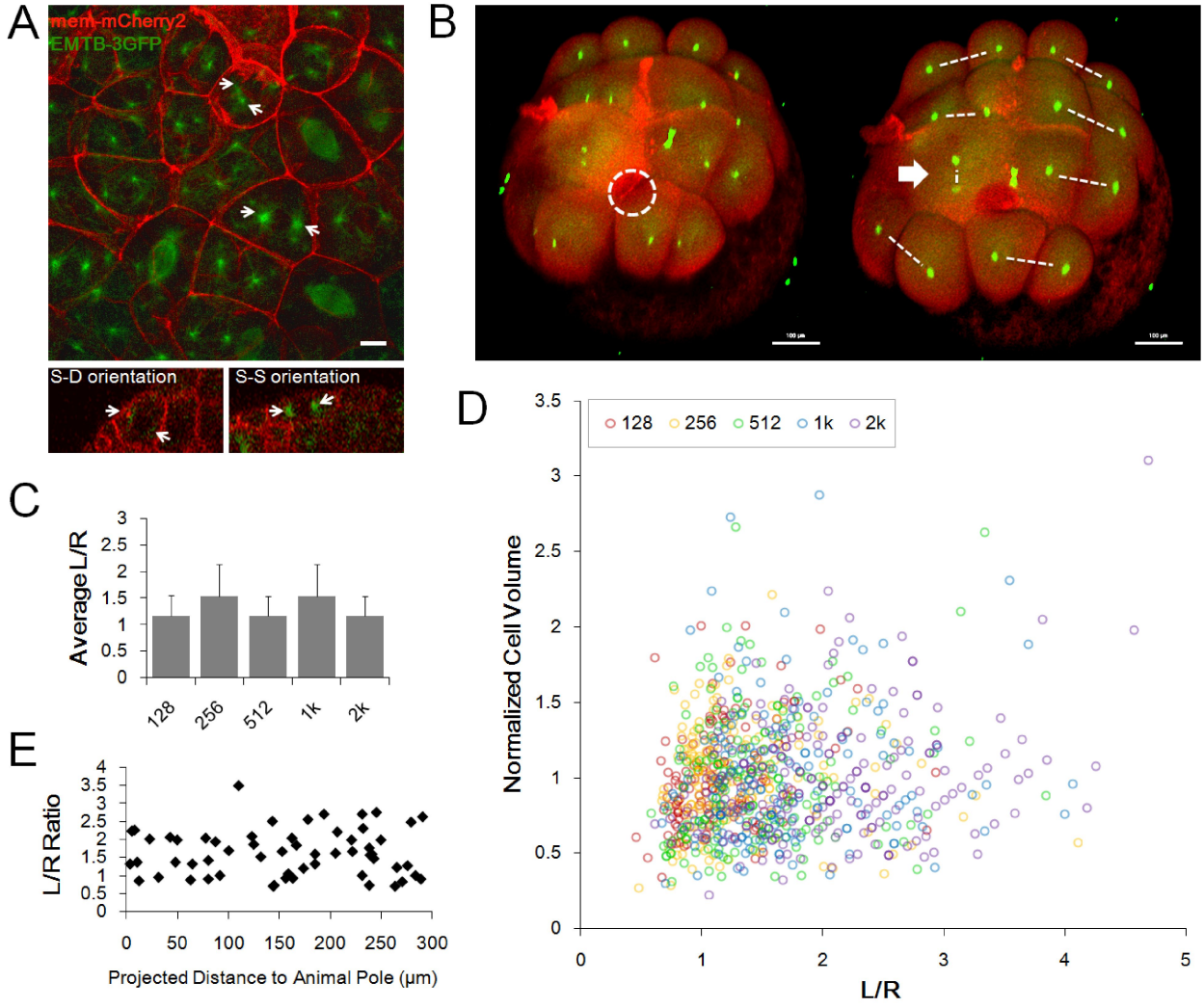


Figure 4.6. Division orientations and geometry models of cell shape distribution

(A) Centrioles and spindles highlighted by *tg(actb2:EMTB-3GFP)* embryos (Wuhr et al., 2010). Arrows: centrioles. Scale Bar: 10 µm.

(B) Changed division orientation by oil drop injection. Dashed white circle highlights the injected oil drop inside one of the 8 early blastomeres. This blastomere (Arrow) bulges at the top surface. Consequentially it divides perpendicularly to the surface generating an aberrant S-D division. Dashed lines highlight division orientation. Scale Bars: 100µm.

(C) Morphogenesis predicted by a pure geometry-division model, applying the division rule

Figure 4.6 (*Continued*) Th=1.25 and using 128-cell stage cell shape distribution as initial input. The result is oscillating average L/R ratios instead of observed gradual stretching.

(D) Correlation between cell volume and L/R ratio at all time points. Cell volume does not predict L/R ratio of the cell therefore the volume distribution is unlikely underlying the shape distribution.

(E) Spatial distribution of shapes. Cells were randomly sampled from different distances of the animal pole (projected distance ignores the surface curvature). Cells of different L/R ratios are found at different locations therefore spatial heterogeneity of shapes is unlikely underlying the shape distribution.

Together, our results have interesting implications: since the shape of a single cell predicts its division orientation, the cell morphology distribution (Figure 4.1E) becomes predictive of the fraction of S-D divisions of the population, which further determines the next N_S. This suggests that not only the average L/R, but also the distribution of L/R, is an important property governing epithelial morphogenesis. To find the cause of the observed evolution of cell shape distributions, we first explored simple mechanisms that do not account for cell-cell interactions. For example, pure geometrical partitioning from an initial distribution (Figure 4.6C), shape variation from unevenness of cell volumes (Figure 4.6D), or spatial heterogeneity across the embryo (Figure 4.6E). These geometrical models do not explain the observed distributions, indicating mechanical interactions between cells need to be considered.

Mechanical interactions define the distribution of cell shapes

To understand why cell shapes form a wide distribution of a certain non-symmetric shape, we considered the local factors that affect single cell morphology and their interactions on the population level. Cell shape is controlled by a variety of mechanical forces, from global, local and internal sources (Figure 4.7A, Heisenberg and Bellaiche, 2013). Tissue surface area A, as a geometrical property of the whole layer, influences single cells as the forces that constrain A are transduced through cell-cell junctions. Cells interact with neighbors using adhesion molecules which tend to increase cell-cell contact surface and deform the cell. In the absence of cell-cell adhesion, the cell tends toward a spherical morphology due to internal cortical tension (Adams et al., 1998; Kafer et al., 2007; Manning et al., 2010; Figure S4.7A). The observed morphological distribution likely represents a force-balanced state for the population, which is disrupted by cortical force changes during divisions and then rebalanced as daughter cells reform stable shapes until the next division.

We modeled the cell shapes with an energy function by assuming each cell morphology has an energy value and the cells tend to minimize the energy of the whole population. The surface area constraint A is implemented as a fixed sum of free surface area of individual cells (Figure 4.7A). In addition, each cell reduces energy by reducing total cell surface area (cortical tension) or forming contact surfaces with other cells (adhesion) - competing tendencies that result in a defined cell energy function with a single minimum (SI text.3, Figure 4.8B, Manning et al., 2010). While cells overall change shapes to move towards a lower energy level, there is certain randomness in the shape changes such that some cells will be at a higher energy level than others

at any time of observation. The probability of a certain cell shape to be observed is then a function of its own energy and its energy impact on the population (Figure 4.7B). In implementation, this model uses two parameters describing the mechanical interactions: **γ**, which represents the ratio of adhesion over cortical tension per unit cell contact surface area change, and **T** (analogous to temperature in a thermodynamic approach), which represents the sensitivity of the system to energy change and is the noise property of γ. With A and N_S as inputs, the model successfully recaptures the shapes of observed L/R distributions at all time points (Figures 4.7C-D, SI text.4). We then took one step further to implement the division rule (Th, Sh) into the model (Figure 4.8C) to compute the fraction of S-D divisions given a modeled L/R distribution and use the predicted N_S to determine the next L/R distribution, leading to a full interplay model (SI text.5). This model satisfactorily reproduces the morphogenesis process, albeit being less faithful at the final time point of 2k cell stage (Figure 4.7E) compared to using measured N_S as input.

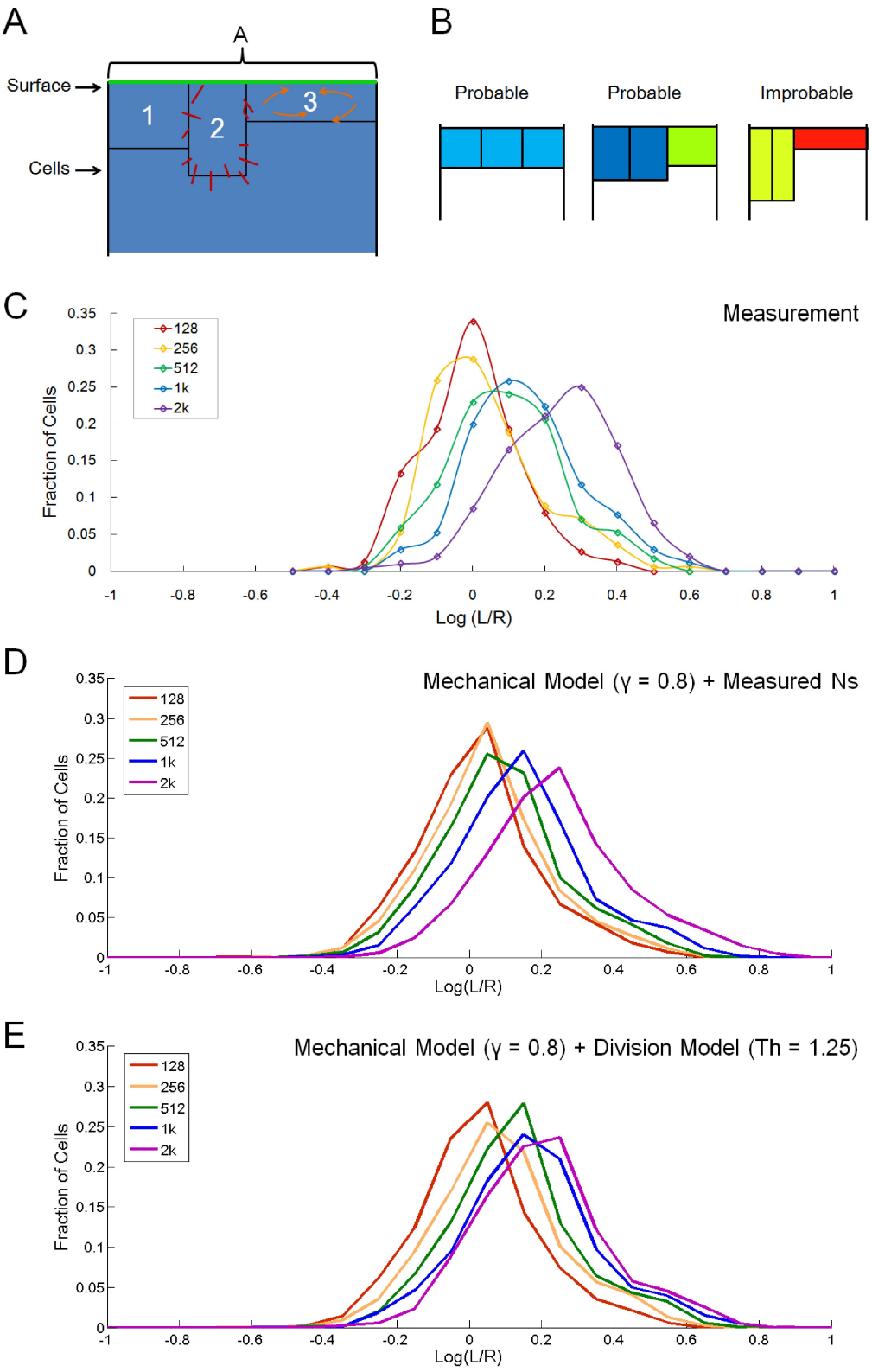


Figure 4.7. Mechanical modeling recaptures cell shape distribution and dynamics

Figure 4.7 (*Continued*) (A) Schematic illustration of a surface configuration. Cells are constrained to cover a surface area of A (green line). Cells form adhesion junctions with other cells (red lines in cell#2) and cortical tension pulls cells to become more isometric (orange arrows in cell#3). The illustration is shown in 2D but it was modeled in 3D. See also Figures 4.8A-B, SI text.3.

(B) Different surface configurations may appear and the probability is determined by the overall energy level (represented by blue to red colors) of the configuration. A shape with high energy may be more likely to appear if its occurrence reduces the energy of other cells. See also Figures 4.8B-C, SI text.4.

(C) Data in Figure 4.1E re-plotted with fraction/log axes to facilitate comparison with simulations.

(D) Simulation results using the mechanical model and measured N_S. See also SI text.4.

(E) Simulation results incorporating both mechanical and division models (Interplay model). See also SI text.5.

Together, these results support the model that mechanical interactions and the division rule define the cell shape distribution and its evolution over time, through a step-by-step feedback cycle between the shape distribution and N_S to ultimately create a simple epithelium of a characteristic cell morphology. The dynamics of surface morphogenesis therefore is dependent on Th and γ. Analyzing their properties will be important for understanding how this system can produce different morphological outputs.

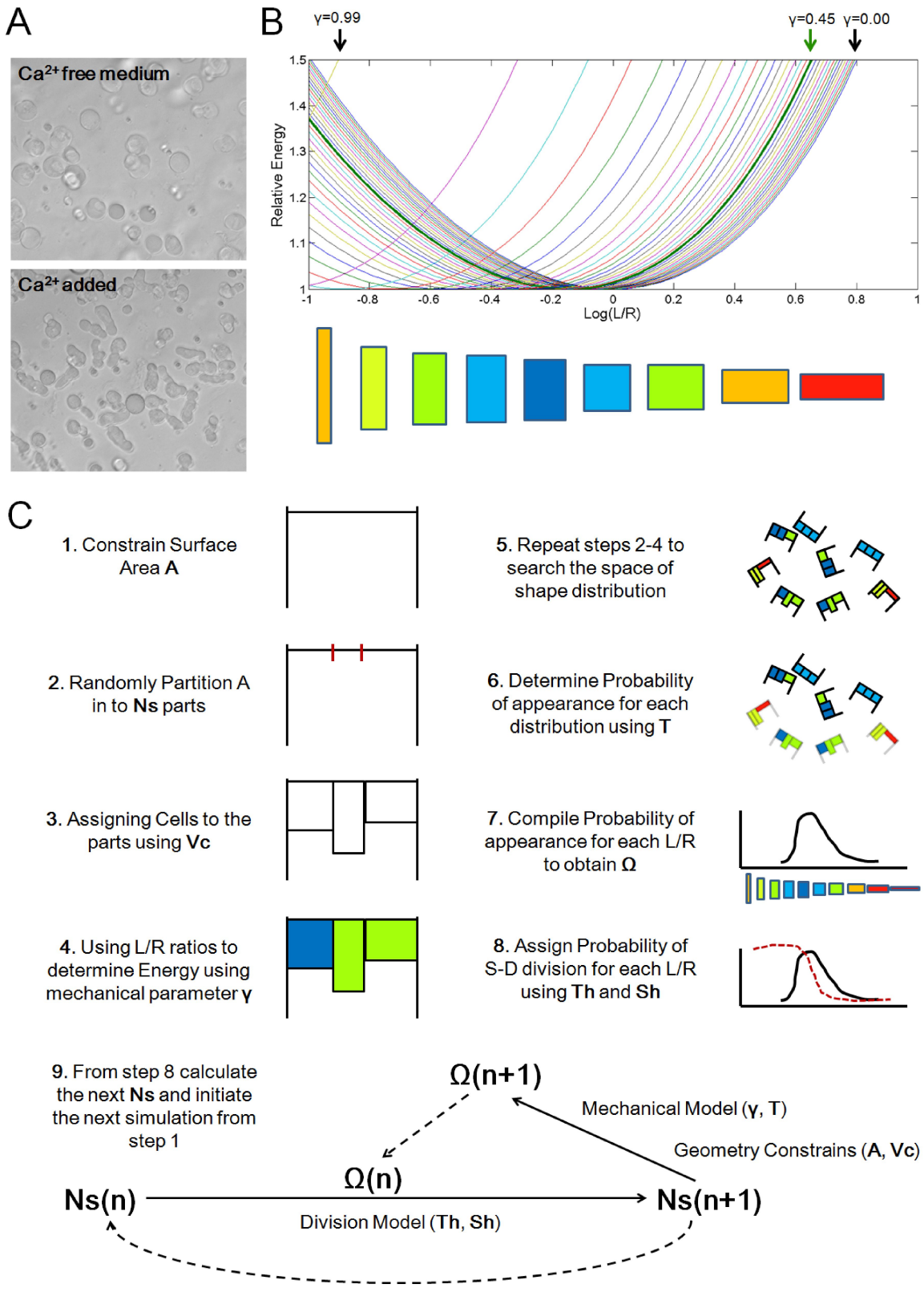


Figure 4.8. Mechanical modeling of cell shape distributions

Figure 4.8 (*Continued*) (A) Dissociated cells in Ca^{2+} free medium round up as spheres under cortical tension, adding Ca^{2+} allows them to recover adhesion to attach to dish surface or other cells and deform from sphere.

(B) The relationship between cell shape and cell energy as a function of parameter γ. Relative energy is computed by normalization to the minimal energy. Depending on the value of γ the low energy cell shape is different. An example (γ=0.45) of corresponding cell shapes and energy (represented by blue to red colors) is shown below the graph. See also SI text.3.

(C) Full streamline of interplay model of surface morphogenesis. See also SI texts.4-5.

Interplay model suggests a general link between morphogenesis and cell division

To understand how the key parameters of the division rule (Th) and the mechanical rule (γ) in the interplay model affect the morphogenesis process, we performed *in silico* experiments to alter their values. Interestingly, the change of Th produces a full spectrum of resulting epithelial morphologies from squamous to columnar (Figure 4.9A) and corresponding N_S changes (Figure 4.9B). Parameter Sh defines the sharpness of the threshold switch. As expected, a low Sh will bring the fraction of S-D division closer to 0.5 with the cell shape distribution unchanged (Figures 4.10A-B). We are unable to specifically perturb Th *in vivo* to validate these *in silico* results. However, since our interplay model does not depend on assumptions restricted to zebrafish EVL except it being a monolayer of surface cells, we reasoned that the link between cell division control and collective morphogenesis may be general. To test this, we applied our model to frog and sea urchin embryos (Wray 1997; Chalmers et al., 2003; Wuhr et al, 2008),

where the tissue surface area A likely also remains constant and cell volume V_C also roughly halves during each division, but the division rule is quite different (i.e. the main difference from fish lies in the parameter Th). By inputting corresponding initial conditions and an altered Th value into our model (SI text.6), surprisingly, we are able to predict the surface morphogenesis as well as division patterns of these systems (Figures 4.9E-F, 4.10E-F). These data suggest that division orientation control as a function of cell shape plays a key role in generating distinct epithelial morphologies in general.

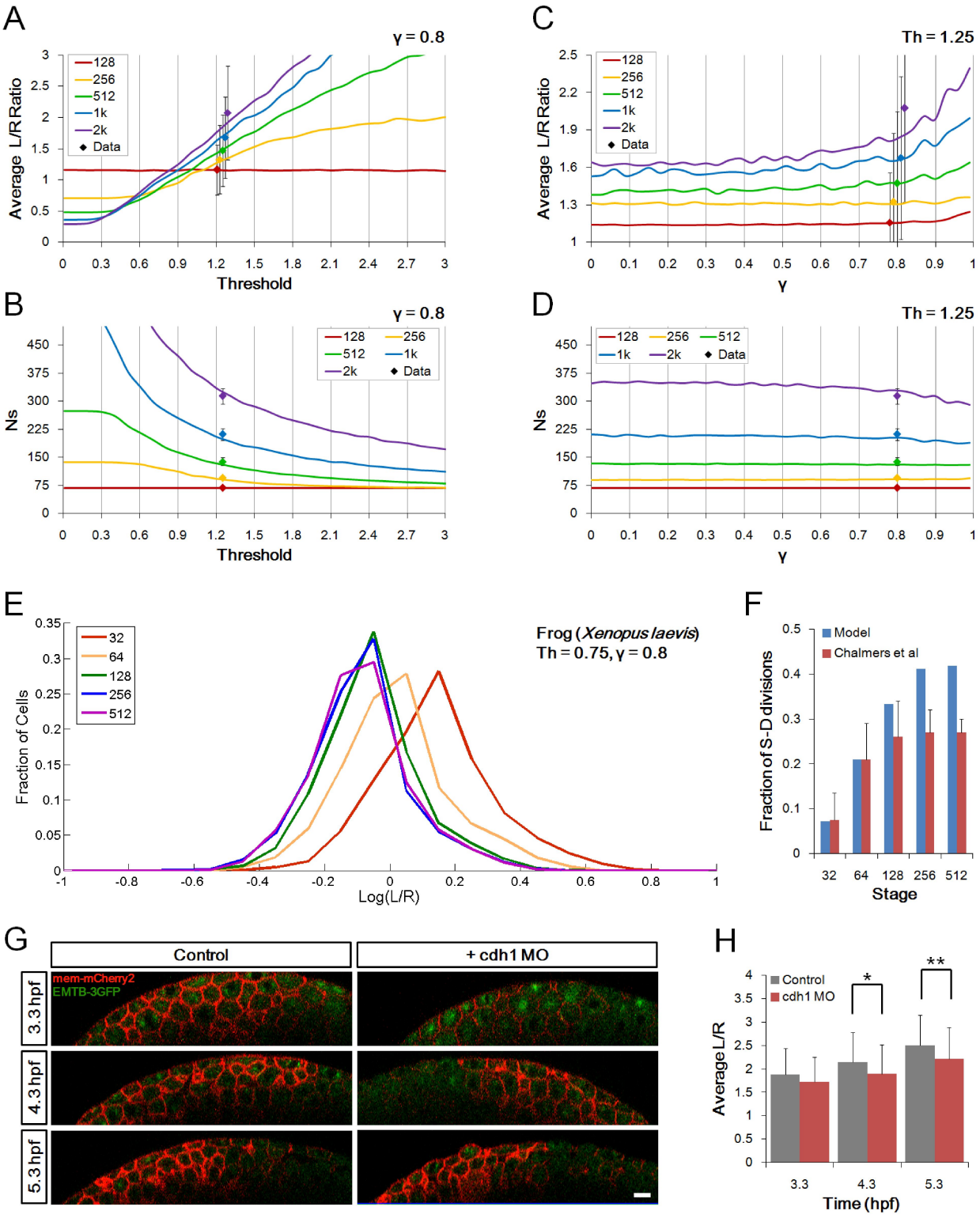


Figure 4.9. Effect of altering Th and γ in surface morphogenesis

(A) Interplay model simulation of average L/R using variable Th values under a fixed γ=0.8.

Figure 4.9 (*Continued*) Measurement data points are plotted on Th=1.25 with slight shifts between points to better visualize error bars (s.d.).

(B) Same simulation as (A) plotting N_S dynamics.

(C) Interplay model simulation of average L/R using variable γ values under a fixed Th=1.25. Measurement data points are plotted on γ=0.8 with slight shifts between points to better visualize error bars (s.d.).

(D) Same simulation as (C) plotting N_S dynamics.

(E) Simulation of frog embryo surface morphogenesis using an altered Th value and different initial conditions. See also SI text.6.

(F) Prediction of S-D division ratio by model in (E) and comparison with measurements by Chalmers et al., 2003 (Data approximated from Fig.5B of this reference).

(G) Timelapse imaging data of control and cdh1 MO injected embryos. Scale bars: 20µm.

(H) Reduced flattening of cdh1 MO injected embryos, n=60 cells from 3 embryos at each time point were measured for both groups. Error bars are s.d. At 3.3h p=0.11; *p=0.03; **p=0.02 (two tail t-test).

The increase of γ increases average L/R (Figure 4.9C) and decreases N_S (Figure 4.9D). Parameter T determines the variance of cell shape distributions. As expected, a high T allows more high energy cell shapes to appear which also affects the fraction of S-D divisions (Figures 4.10C-D). Interestingly, the effect of γ change is more apparent in later stages than in earlier stages (Figure 4.9C). This suggests that if cell adhesion increases over time (effectively

increasing γ), the model prediction will be more accurate compared to Figure 4.7E (where γ is constantly 0.8) as the 2k time data point will shift to the right onto the curve in Figure 4.9C. Consistent with this prediction, the adhesion molecule E-cadherin (cdh1) expression increases in the surface cells over this time (Kane et al., 2005). Conversely, the model predicts that reducing adhesion (reducing γ) will delay flattening, a phenotype that will become more evident in later stages. To test this prediction, we injected cdh1 morpholino (MO) to reduce adhesion. Indeed, while the surface of injected embryos is not significantly different in flattening by 2k stage ($p=0.11$), the difference becomes apparent in the next 2 hours (Figures 4.9G-H, *$p=0.03$; **$p=0.02$) as the model suggested. Furthermore, the morpholino injected embryos are frequently found to have cell extrusions (11/15 vs. control 0/12), a possible indicator of excessive surface cells (larger N_S) (Eisenhoffer et al., 2012) also predicted by the model (Figure 4.9D).

Together, these results show that different parameters influence surface morphogenesis differentially thus providing a basis of generating diversity. Importantly, system properties such as Th and γ are local and can thus be changed by molecular evolution and regulated by gene expression changes during patterning. Our model makes non-intuitive yet testable and quantitative predictions and is therefore useful for unraveling the complex interactions of factors that participate in the morphogenesis of simple epithelia.

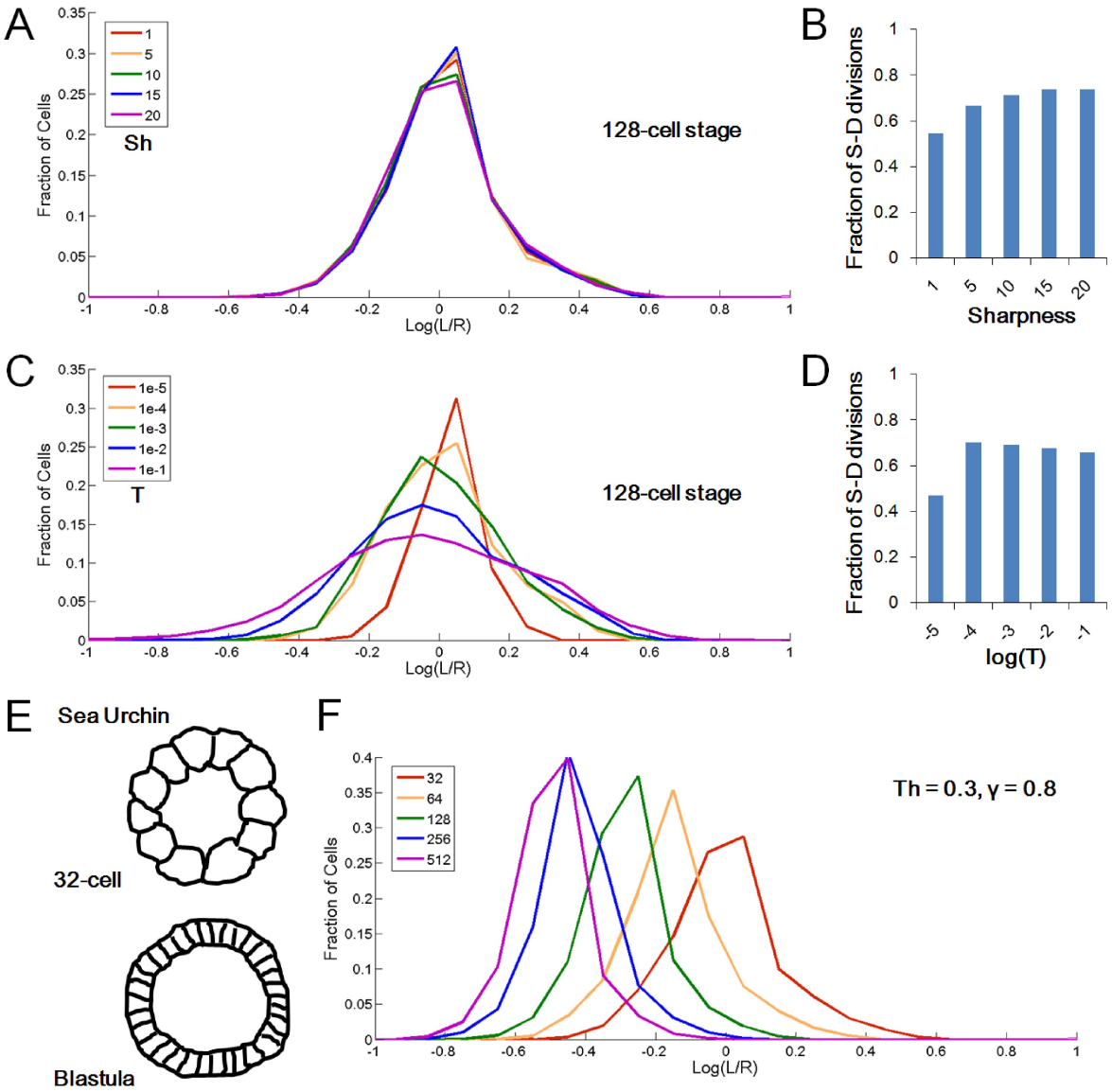


Figure 4.10. Effects of altering Sh and T and sea urchin embryo surface modeling

(A) Interplay model simulation of average L/R using variable Sh values. Other parameters: Th=1.25,γ=0.8,T=1e-4. Inputs at 128-cell stage were used.

(B) Same simulation as (A) plotting Fraction of S-D divisions under different Sh values.

(C) Interplay model simulation of average L/R using variable T values. Other parameters: γ=0.8,Th=1.25, Sh=10. Inputs at 128-cell stage were used.

(D) Same simulation as (C) plotting Fraction of S-D divisions under different T values.

Figure 4.10 (*Continued*) (E) Schematic of sea urchin embryo cross-section surface cell shapes. Adapted from photos in Wray 1997.

(F) Simulation of sea urchin embryo surface morphogenesis using an altered Th value and different initial conditions. See also SI text.6.

DISCUSSION

The relationship between surface morphogenesis and EVL differentiation

EVL is an important model for studying epithelial differentiation (Fukazawa et al., 2010). The cues that trigger the activation of an EVL-specific genetic program remains unclear. One possibility is that the flattening of the surface cells itself is the cue (as proposed by Manning et al. 2010, in a communication with Krens et al. 2011). We have shown that the fate commitment (lineage separation) of surface cells to EVL is after the 12th cell cycle, consistent with previous lineage tracing and transplantation experiments (Ho 1992; Kane et al., 1992). In addition, the timing of EVL reporter onset occurs ~3 hours after apparent flattening that begins before mid-blastula transition. The earliest detectable EVL markers as measured by *in situ* hybridization are not before late blastula stage (Sagerström et al., 2005). The available data thus show that shape changes of surface cells happen much earlier than specification. Together these data support the hypothesis that cells are initially specified to become EVL by the geometric and mechanically induced shape changes described here rather than via a typical molecular trigger. However, an alternative mechanism of asymmetric division is also plausible (Chalmers et al., 2003). In this

scenario, due to the apical-basal polarity of the surface cells, S-D divisions along the lineage may lead to accumulation of important pro-EVL factors in the surface cells, which eventually activate an EVL genetic program. An intriguing observation is that surface cells are known to show Ca^{2+} transients (which can be activated by Wnt/Ca^{2+} signaling) more frequently than deep cells, and the number of transients is associated with cell shape changes (Zhang et al., 2011). What molecular changes follow surface flattening and their relationship to eventual EVL differentiation are not considered in our current model and are important problems for future investigations.

Mechanical links between cell shape and division orientation

Consistent with other studies (Hertwig 1884; Chalmers et al., 2003; Baena-Lopez et al., 2005; Gibson et al., 2011), our results show that the division orientation of surface cells is strongly biased by cell shape, following a "long axis rule" (Hertwig 1884) in general but is not deterministic near the division type switching threshold. Our model predicts that differences in this threshold can produce distinct morphogenetic behaviors. However, how the threshold is determined is unclear. Apparently it is a tunable parameter in nature considering the wide range of division modes in different surface tissues that emerge in the same embryo. To measure its own shape, cells utilize a microtubule network to probe the cortex which in turn generate forces on the nucleus and spindle to determine the division plane (Wuhr et al., 2010; Minc et al., 2011). This mechanism allows intrinsic and extrinsic polarity cues to influence the force balance, which can cause polarized anchoring of centrosomes and/or rotation of the spindle (Galli and van den

Heuvel, 2008; Rebollo et al., 2009; Peyre et al., 2011). Therefore the observed threshold values are likely reflective of how such polarity cues interact with or sometimes override the "default" cell shape/long axis determinant (Gillies and Cabernard, 2011). For example, the dominance of S-S divisions in sea urchin embryos might be a result of the strong apical anchoring of the cells to the hyaline layer (Wolpert and Gustafson, 1961). Understanding how the apical basal polarity of surface cells influences the microtubule network mechanically may reveal the basis of threshold values and explain the observed differences between systems such as fish, frog and sea urchin.

Producing tissue behavior using local shape control and divisions

Our model shows that the morphogenesis towards surface cell shapes of high and low L/R values increases the mechanical energy of the tissue, implying these outcomes require considerable forces to maintain and conversely the counterforces may produce desired morphogenetic behaviors. For example, buckling or invagination may occur to increase surface area when cells keep making S-S divisions or continue to grow, as highly crowded columnar cells exert a pushing force against the geometrical constraining borders (Eiraku et al., 2011; Sato and Clevers, 2013). On the other hand, a highly stretched surface layer generates strong pulling forces on its edge, which may lead to further spreading or shrinking of the surface, depending on conditions. An example of this is the epiboly of zebrafish embryos, where the spreading EVL requires strong connections to the yolk syncytial layer (YSL) that pull the EVL cells (Solnica-Krezel and Driever, 1994; Koppen et al., 2006; Behrndt et al., 2012). When YSL formation is disrupted the

surface shrinks back causing a yolk burst (Xu et al., 2012). Currently, the mechanism that initiates epiboly movements is unclear. Our work provides an intriguing possibility that the flattening of surface cells in early stages prior to surface area expansion might generate the forces that initiate the doming (bulging of yolk into the cells before epiboly) and subsequent epiboly movements.

Robustness of morphogenesis from the interplay between shape and division

Developmental processes exhibit robustness in reaching "canalized" goals (Waddington 1942). Such robustness is often achieved through feedback interactions (Alon 2007) that allow genetic regulatory networks or cells to "steer" the dynamic process to the correct track in the face of noise, errors, and changing environments. Our model suggests that the feedback to N_S by the cell shape distribution underlies the robustness of surface flattening. For example, surface cell loss (which lowers N_S) makes the neighbors of the lost cell stretch more to cover its previously occupied surface area, biasing their next division choice towards S-S, thus compensating the loss (by increasing N_S). Our model also suggests that given a constant A and a halving V_C the cell shape distribution will move towards a final steady state defined by Th and γ, independent of the initial distribution. These properties not only provide robustness but also render global, tissue-level morphogenetic goals under the control of local, cell behavioral regulations that can be exploited by molecular changes during development and evolution.

EXPERIMENTAL PROCEDURES

Zebrafish strains and maintenance

See Extended Experimental Procedures for protocols, sources and references for transgenic strains used in this study. All fish related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Microinjections of mRNAs and morpholinos (MO)

1-2 cell stage embryos were injected (Nanoject) 2.3nl 20ng/µl of mem-mCherry2 mRNA with/without 0.1mM cdh1 MO (5-ATCCCACAGTTGTTACACAAGCCAT-3 (Babb and. Marrs, 2004), GeneTools Inc).

Timelapse 2-photon/confocal imaging

1-64 cell stage embryos were dechorionated using sharp tweezers (Dumont 55) and mounted animal pole up (or sideways) into a dorsal mount (Megason 2009). Live imaging was performed using a Zeiss 710 confocal microscope (objectives: Plan-Apochromat 20X 1.0 NA, C-Apochromat 40X 1.2 NA) with a home-made heating chamber maintaining 28ºC. Laser lines 488nm, 514nm, 561nm and 594nm were used. See Figure S4.1, Extended Experimental Procedures for details.

Image Data analysis

Segmentation and tracking were performed using GoFigure2 (www.gofigure2.org) and ACME (Mosaliganti et al., 2012). Cell counting was carried out using ImageJ Cell counter plugin after exporting 3D surface rendering images of the raw datasets from FluoRender (Wan et al., 2009). Measurement of L/R ratios was carried out using ZEN (Carl Zeiss) software 3D distance functionality. Measurements were plotted with Microsoft Excel. See Figure 4.4, Movie S4.2,

Extended Experimental Procedures for more details.

Modeling

Simple geometry models were drawn on paper and calculated in Microsoft Excel. The full interplay model and its simulations were performed using Matlab (Mathworks). See Figures 4.8,4.10, SI texts.1-6 for details.

ACKNOWLEDGEMENTS

We thank D. D'India for fish care, H. Otsuna for movie making, C. Heisenberg for the *tg(krt4:EGFPcaax)* line, P. Keller, M. Wuhr for communications, Megason lab members and T. Mitchison for comments. This work is supported by NIH grants HG004071 and DC010791. F. X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

CHAPTER 4 REFERENCES

Adams, C.L., Chen, Y.T., Smith, S.J., and Nelson, W.J. (1998). Mechanisms of epithelial cell-cell adhesion and cell compaction revealed by high-resolution tracking of E-cadherin-green fluorescent protein. J. Cell Biol. *142*, 1105-1119.

Alon, U. (2007) Network motifs: theory and experimental approaches. Nat. Rev. Genet. *8*,450–461.

Babb, S.G., and Marrs, J.A. (2004). E-cadherin regulates cell movements and tissue formation in early zebrafish embryos. Dev. Dyn. *230*, 263-277.

Baena-Lopez, L.A., Baonza, A., and Garcia-Bellido, A. (2005). The orientation of cell divisions determines the shape of Drosophila organs. Curr. Biol. *15,* 1640-1644.

Behrndt, M., Salbreux, G., Campinho, P., Hauschild, R., Oswald, F., Roensch, J., Grill, S.W., and Heisenberg, C.P. (2012). Forces driving epithelial spreading in zebrafish gastrulation. Science *338,* 257-260.

Chalmers, A., Strauss, B., and Papalopulu, N. (2003). Oriented cell divisions asymmetrically segregate aPKC and generate cell fate diversity in the early Xenopus embryo. Development *130,* 2657-2668.

Ciruna, B., Jenny, A., Lee, D., Mlodzik, M., and Schier, A. (2006). Planar cell polarity signalling couples cell division and morphogenesis during neurulation. Nature *439,* 220-224.

Deramaudt, T., and Rustgi, A.K. (2005). Mutant KRAS in the initiation of pancreatic cancer. Biochim. Biophys. Acta *1756,* 97-101.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature *472,* 51-56.

Eisenhoffer, G.T., Loftus, P.D., Yoshigi, M., Otsuna, H., Chien, C.B., Morcos, P.A., and Rosenblatt, J. (2012). Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. Nature *484,* 546-549.

Fernandes, I., Chanut-Delalande, H., Ferrer, P., Latapie, Y., Waltzer, L., Affolter, M., Payre, F., and Plaza, S. (2010). Zona pellucida domain proteins remodel the apical compartment for localized cell shape changes. Dev. Cell. *18,* 64-76.

Fernandez, R., Das, P., Mirabet, V., Moscardi, E., Traas, J., Verdeil, J., Malandain, G., and Godin, C. (2010). Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. Nat. Methods *7,* 547-553.

Fukazawa, C., Santiago, C., Park, K., Deery, W., Gomez de la Torre Canny,Sol, Holterhoff, C., and Wagner, D. (2010). poky/chuk/ikk1 is required for differentiation of the zebrafish embryonic epidermis. Dev. Biol. *346,* 272-283.

Galli, M., and van den Heuvel, S. (2008). Determination of the cleavage plane in early C. elegans embryos. Annu. Rev. Genet. *42,* 389-411.

Gibson, W., Veldhuis, J., Rubinstein, B., Cartwright, H., Perrimon, N., Brodland, G., Nagpal, R., and Gibson, M. (2011). Control of the mitotic cleavage plane by local epithelial topology. Cell *144,* 427-438.

Gillies, T., and Cabernard, C. (2011). Cell division orientation in animals. Current Biology *21,* 609.

Guillot, C., and Lecuit, T. (2013). Mechanics of epithelial tissue homeostasis and morphogenesis. Science *340,* 1185-1189.

Grill, S.W., and Hyman, A.A. (2005). Spindle positioning by cortical pulling forces. Dev. Cell. *8,* 461-465.

Heisenberg, C., and Bellaiche, Y. (2013). Forces in tissue morphogenesis and patterning. Cell *153,* 948-962.

Hertwig, O. (1884). Das Problem der Befruchtung une der Isotropie des Eies, eine Theory der Vererbung (Jenaische Zeitschrist)

Ho, R.K. (1992). Cell movements and cell fate during zebrafish gastrulation. Dev. Suppl. 65-73.

Kafer, J., Hayashi, T., Maree, A.F., Carthew, R.W., and Graner, F. (2007). Cell adhesion and cortex contractility determine cell patterning in the Drosophila retina. Proc. Natl. Acad. Sci. U. S. A. *104,* 18549-18554.

Kane, D.A., Warga, R.M., and Kimmel, C.B. (1992). Mitotic domains in the early embryo of the zebrafish. Nature *360,* 735-737.

Kane, D.A., McFarland, K.N., and Warga, R.M. (2005). Mutations in half baked/E-cadherin block cell behaviors that are necessary for teleost epiboly. Development *132,* 1105-1116.

Keller, P.J. (2013). Imaging morphogenesis: technological advances and biological insights. Science *340,* 1234168.

Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H. (2008). Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science *322,* 1065-1069.

Keller, P.J., Schmidt, A.D., Santella, A., Khairy, K., Bao, Z., Wittbrodt, J., and Stelzer, E.H. (2010). Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. Nat. Methods *7,* 637-642.

Kimmel, C.B., Ballard, W., Kimmel, S., Ullmann, B., and Schilling, T. (1995). Stages of embryonic development of the zebrafish. Dev. Dyn. *203,* 253-310.

Kimmel, C.B., and Law, R.D. (1985). Cell lineage of zebrafish blastomeres. I. Cleavage pattern and cytoplasmic bridges between cells. Dev. Biol. *108,* 78-85.

Koppen, M., Fernandez, B.G., Carvalho, L., Jacinto, A., and Heisenberg, C.P. (2006). Coordinated cell-shape changes control epithelial movement in zebrafish and Drosophila. Development *133,* 2671-2681.

Krens, S.F., Mollmert, S., and Heisenberg, C.P. (2011). Enveloping cell-layer differentiation at the surface of zebrafish germ-layer tissue explants. Proc. Natl. Acad. Sci. U. S. A. *108,* E9-10; author reply E11.

Krupinski, P., Chickarmane, V., and Peterson, C. (2011). Simulating the mammalian blastocyst--molecular and mechanical interactions pattern the embryo. PLoS Comput. Biol. *7,* e1001128.

Kunda, P., and Baum, B. (2009). The actin cytoskeleton in spindle assembly and positioning. Trends Cell Biol. *19,* 174-179.

Lecuit, T., and Lenne, P. (2007). Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. Nat. Rev. Mol. Cell Biol. *8,* 633-644.

Manning, M.L., Foty, R.A., Steinberg, M.S., and Schoetz, E.M. (2010). Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. Proc. Natl. Acad. Sci. U. S. A. *107,* 12517-12522.

Martin, A.C., Gelbart, M., Fernandez-Gonzalez, R., Kaschube, M., and Wieschaus, E.F. (2010). Integration of contractile forces during tissue invagination. J. Cell Biol. *188,* 735-749.

McMahon, A., Supatto, W., Fraser, S., and Stathopoulos, A. (2008). Dynamic analyses of Drosophila gastrulation provide insights into collective cell migration. Science *322,* 1546-1550.

Megason, S.G. (2009). In toto imaging of embryogenesis with confocal time-lapse microscopy. Methods Mol. Biol. *546,* 317-332.

Megason, S.G, and Fraser, S.E (2003). Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. Mech. Dev. *120,* 1407-1420.

Megason, S.G., Srinivas, S., Dickinson, M.E., and Hadjantonakis, A.K. (2011). Microscopy to mechanism across the scales of development. Curr. Opin. Genet. Dev. *21,* 519-522.

Minc, N., Burgess, D., and Chang, F. (2011). Influence of cell geometry on division-plane positioning. Cell *144,* 414-426.

Mosaliganti, K.R., Noche, R.R., Xiong, F., Swinburne, I.A., and Megason, S.G. (2012). ACME: automated cell morphology extractor for comprehensive reconstruction of cell membranes. PLoS Comput. Biol. *8(12),* e1002780.

Nelson, C., Jean, R., Tan, J., Liu, W., Sniadecki, N., Spector, A., and Chen, C. (2005). Emergent patterns of growth controlled by multicellular form and mechanics. Proc. Natl. Acad. Sci. U. S. A. *102,* 11594-11599.

Odell, G., Oster, G., Alberch, P., and Burnside, B. (1981). The mechanical basis of morphogenesis. I. Epithelial folding and invagination. Dev. Biol. *85,* 446-462.

Olivier, N., Luengo-Oroz, M., Duloquin, L., Faure, E., Savy, T., Veilleux, I., Solinas, X., Débarre, D., Bourgine, P., Santos, A., Peyriéras, N., and Beaurepaire, E. (2010). Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. Science *329,* 967-971.

Osterfield, M., Du, X., Schupbach, T., Wieschaus, E., and Shvartsman, S.Y. (2013). Three-dimensional epithelial morphogenesis in the developing Drosophila egg. Dev. Cell *24,* 400-410.

Peyre, E., Jaouen, F., Saadaoui, M., Haren, L., Merdes, A., Durbec, P., and Morin, X. (2011). A lateral belt of cortical LGN and NuMA guides mitotic spindle movements and planar division in neuroepithelial cells. J. Cell Biol. *193,* 141-154.

Rebollo, E., Roldan, M., and Gonzalez, C. (2009). Spindle alignment is achieved without rotation after the first cell cycle in Drosophila embryonic neuroblasts. Development *136,* 3393-3397.

Sagerström, C., Gammill, L., Veale, R., and Sive, H. (2005). Specification of the enveloping layer and lack of autoneuralization in zebrafish embryonic explants. Dev. Dyn. *232,* 85-97.

Salazar-Ciudad, I., Jernvall, J., and Newman, S.A. (2003). Mechanisms of pattern formation in development and evolution. Development *130,* 2027-2037.

Sato, T., and Clevers, H. (2013). Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science *340,* 1190-1194.

Solnica-Krezel, L., and Driever, W. (1994). Microtubule arrays of the zebrafish yolk cell: organization and function during epiboly. Development *120,* 2443-2455.

Tamulonis, C., Postma, M., Marlow, H., Magie, C., de Jong, J., and Kaandorp, J. (2011). A cell-based model of Nematostella vectensis gastrulation including bottle cell formation, invagination and zippering. Dev. Biol. *351,* 217-228.

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139,* 871-890.

Waddington, C.H. (1942). The canalization of development and the inheritance of acquired characteristics. Nature *150,* 563-5.

Wan, Y., Otsuna, H., Chien, C.B., and Hansen, C. (2009). An interactive visualization tool for multi-channel confocal microscopy data in neurobiology research. IEEE Trans. Vis. Comput. Graph. *15,* 1489-1496.

Wang, Y., Khan, Z., Kaschube, M., and Wieschaus, E. (2012). Differential positioning of adherens junctions is associated with initiation of epithelial folding. Nature *484,* 390-393.

Wolpert, L., and Gustafson, T. (1961). Studies on the cellular basis of morphogenesis of the sea urchin embryo. The formation of the blastula. Exp. Cell Res. *25,* 374-382.

Wray, G.A. (1997). Echinoderms. *In* S. F. Gilbert and A. Raunio (eds.), *Embryology: Constructing the Organism.* Sinauer Associates, Sunderland, MA. pp. 309–329.

Wuhr, M., Chen, Y., Dumont, S., Groen, A., Needleman, D., Salic, A., and Mitchison, T. (2008). Evidence for an upper limit to mitotic spindle length. Current Biology *18,* 1256-1261.

Wuhr, M., Tan, E., Parker, S., Detrich, H., and Mitchison, T. (2010). A model for cleavage plane determination in early amphibian and fish embryos. Current Biology *20,* 2040-2045.

Xiong, F., Tentner, A.R., Huang, P., Gelas, A., Mosaliganti, K.R., Souhait, L., Rannou, N., Swinburne, I.A., Obholzer, N.D., Cowgill, P.D., Schier, A.F., and Megason, S.G. (2013). Specified neural progenitors sort to form sharp domains after noisy Shh signaling. Cell *153,* 550-561.

Xu, C., Fan, Z.P., Muller, P., Fogley, R., DiBiase, A., Trompouki, E., Unternaehrer, J., Xiong, F., Torregroza, I., Evans, T., *et al.* (2012). Nanog-like regulates endoderm formation through the Mxtx2-Nodal pathway. Dev. Cell. *22,* 625-638.

Zhang, J., Webb, S., Ma, L., Chan, C., and Miller, A. (2011). Necessary role for intracellular Ca^{2+} transients in initiating the apical-basolateral thinning of enveloping layer cells during the early blastula period of zebrafish development. Dev. Growth Differ. *53*, 679-696.

MOVIE CAPTIONS

Movie.S4.1, supporting Figure 4.1

In toto imaging of EVL formation. Cyan: cell membrane, Yellow: cell nuclei. Movie is a top view 3D rendering of the raw timelapse dataset starting at 2-4 cell stage. The temporal resolution between each frame is 2 minutes. Time annotation is hh/mm post fertilization. Scale bar: 100µm.

Movie.S4.2, supporting Figure 4.3

3D membrane (Green) and nuclear (Red) segmentations of surface and deep cells rendered using ACME (Mosaliganti et al., 2012) and GoFigure2 (www.gofigure2.org). Individual cells can be selected (highlighted as red) to extract geometrical information. This method provides the most accurate V_C value which we used to validate whole embryo and L^2R estimations.

SI TEXTS

SI text.1. The average L/R ratio and the estimation by $L/R=(A/N_S)^{1.5}/V_C$ (eq.1)

For each single cell, our cube shape simplification provides a simple relationship between its L/R ratio and its surface area and volume, which is:

$$L/R = \frac{A_C{}^{1.5}}{V_C}$$

This relationship holds for the population under an ideal configuration where all cells are the same, which is:

$$< L/R >= \frac{(A/N_S)^{1.5}}{V_C} \quad \text{(eq.1)}$$

The measured average L/R ratio is calculated as:

$$< L/R >= \frac{1}{N_S}\left(\sum_1^{N_S} \frac{L_i}{R_i}\right) = \frac{1}{N_S}\left(\sum_1^{N_S} \frac{Ac_i^{1.5}}{Vc_i}\right) \quad \text{(eq.2)}$$

To perform the quantitative correlations in Figure 4.3D, it is important to consider the possible differences of average L/R values obtained from eq.1 and eq.2. The measurements show that L_i/R_i forms an asymmetric distribution but is close to a normal distribution in log plots. This means that the actual distribution has a long tail on the right side of the average point and a short but higher peak on the left side (i.e. a smaller number of high L/R cells contribute to the increase of the average vs. a large number of cells with slightly below average L/R contribute to the decrease). This would cause the estimation value using eq.1 to be smaller than that of eq.2, as the ideal uniform cell shape assumption of eq.1 will introduce more errors for L/R values further from the average, thus underestimate the weight of high L/R values. This was the observed case in Figure 4.3D. The measured mean, median, eq.1 estimation and global estimation values are list in the following table:

Table S4.1 L/R values of surface cells by different measures.

Surface L/R values	**128**	**256**	**512**	**1k**	**2k**
mean	1.16	1.32	1.47	1.68	2.08
median	1.08	1.16	1.36	1.51	2.00
eq.1	1.15	1.27	1.43	1.61	2.03
whole embryo estimation	1.14	1.40	1.58	1.68	1.86

SI text.2. Rigid cell model

A useful boundary condition of N_S changes that can be used to compare with measurement in order to determine if crowding or flattening would appear is the rigid cell model. In this model cells are rigid spheres that cannot change morphology. Using the surface area constrain and volume change rules (which becomes half every division in this case), the model can estimate how many rigid cells are needed to cover the surface of the embryo. An observed trend lower than this would imply flattening and a higher trend implies crowding, respectively.

Using basic geometry, we have a big sphere assembled from small spheres, the number of surface spheres (N_S) is given by:

$$Ns = N - \frac{V_{deep}}{Vc} = N - \rho \frac{\frac{4\pi}{3}(r-2r_c)^3}{\frac{4\pi}{3}{r_c}^3} = N - \rho(\frac{r}{r_c} - 2)^3 \qquad \text{(eq.4)}$$

where ρ is the packing density which is 0.64 for spheres. r/r_c is given by:

$$\frac{r}{r_c} = (\frac{V_{sphere}}{Vc})^{1/3} = (\frac{V_{material}/\rho}{V_{material}/N})^{1/3} = (\frac{N}{\rho})^{1/3} \qquad \text{(eq.5)}$$

Therefore, the number of spheres on the surface is a function of N, total number of spheres. Consider the fish embryo as a hemisphere, the formula for N_S will then be:

$$Ns = N - \frac{\rho}{2}((\frac{2N}{\rho})^{1/3} - 2)^3 \qquad \text{(eq.6)}$$

Eq.6 provides the sequence of N_S plotted in Figure 4.3C.

SI text.3. Single cell energy model incorporating adhesion and cortical tension

To understand how L/R ratio is influenced by cell mechanics, we consider adhesion and cortical tension for each single cell. An energy (Ec_i) that is a function of the cell's volume (Vc_i) and morphology (again L/R) is built as follows: More cell-cell contact area reduces energy, less total

surface area reduces energy, therefore (see also Manning et al., 2010):

$$E_C = \alpha \cdot (2L^2 + 4LR) - \beta \cdot (L^2 + 4LR) \quad \text{(eq.7)}$$

in which α is the cortical energy coefficient per surface area change and β is the adhesion coefficient per surface area change. We will use relative energy for Ec as L and R change a lot as cells change, so we divide eq.7 by α to generate one combined parameter γ=β/α, also we will use L/R and V_C as inputs for this function. These transformations give:

$$E_C{}' = (2L^2 + 4LR) - \gamma \cdot (L^2 + 4LR) = (2-\gamma)L^2 + 4(1-\gamma)LR$$

$$\text{in which: } L = \sqrt[3]{V_C \cdot (L/R)};\ R = \sqrt[3]{V_C/(L/R)^2} \quad \text{(eq.8)}$$

The minimal energy is arrived when:

$$L(Ec_{min}) = \sqrt[3]{2V_C \cdot \left(\frac{1-\gamma}{2-\gamma}\right)} \text{ and } \gamma \leq 1 \quad \text{(eq.9)}$$

When γ>1 it means the reduction of energy by increasing contact surface area will always be preferred over reduction of energy by reducing total surface area, in eq.8 this leads to super squamous or columnar shapes that are not possible in reality. In practice fundamental constrains of plasma membrane area and cell contents prevent extreme shapes. In this study we consider only γ≤1.

By normalizing Ec with Ec_{min} for a range of possible γ values we derive Figure 4.8B.

SI text.4. Shape distribution under constraints - mechanical energy - probability model

The distribution of cell shapes of the surface cell layer is described as a distribution of L/R ratios of all surface cells: Ω. Each Ω must meet the geometrical constraints imposed by the number of

surface cells (N_S, i=1,2,3,...,N_S), the total surface area to cover (A=sum(Ac_i)) and the cell volumes (Vc_i). So Ω is confined by these parameters into a possible Ω space, which is a limited space. Each Ω has an energy (E) that is defined by a sum of energy of single cells as described in SI text.3. There will be one Ω that has the lowest energy E_{min} that is most likely to appear. We assign the probability (P(Ω)) of each Ω to appear as:

$$P(\Omega) = \frac{e^{-(E(\Omega)-E_{min})/T}}{\int e^{-(E-E_{min})/T}} \quad \text{(eq.10)}$$

in which, depending on the noise in this system T, different Ωs may be likely or unlikely to appear.

With above definitions, the probability of finding a cell with a certain L/R value at a random inspection of the population P(L/R) - equivalent to take a measurement of a single cell - is a 2-step discovery: first, find an Ω at P(Ω); second, find an L/R cell in N_S cells in this Ω. This is expressed as:

$$P(L/R) = \int P(\Omega) \cdot \frac{N_{L/R}}{N_S} \quad \text{(eq.11)}$$

Perform this search many times (For detailed implementation see scripts in supplementary files), effectively probing the Ω space with enough coverage, an accurate P(L/R) can be arrived at. This generates the profile of most likely/practical L/R distribution Ω under mechanical energy functions (specified by parameter γ and T), and the N_S, A, V_C constrains. This method produces the shape distributions as seen in the figures. See also Figure 4.6C.

SI text.5. Full interplay model incorporating the division rule

Based on the mechanical model described in SI text.4. The division rule which suggests that Ω predicts what happens during the next round of divisions. The fraction of cells under certain L/R=Th will more likely undergo S-D division, others will more likely undergo S-S division, which could be quantitatively simulated using the Hill function (specified by Th and Sh parameters) to assign a probability of S-D division for each cell:

$$P_{SD} = \frac{Th^{Sh}}{Th^{Sh}+(L/R)^{Sh}} \quad \text{(eq.12)}$$

On the population level, we have $N_S(n+1)= N_S(n)(2-\Phi)$, in which Φ is the fraction of S-D divisions at nth time. Using eq.11 and eq.12 we have (For detailed implementation see scripts in supplementary files):

$$\Phi = \frac{\sum P(L/R)\times P_{SD}}{N_S} \quad \text{(eq.13)}$$

This provides us with $N_S(n+1)$ which is to be used for determining Ω(n+1) thus forming a complete stepwise interplay model that does not depend on N_S measurements to simulate morphogenesis, instead it is capable of predicting N_S changes. See also Figure 4.6C.

SI text.6. Parameter analysis and application to other systems

The full interplay model described in SI text.5 allows us to probe parameter properties, for example Th, Sh, γ and T. Loop scripts were used to explore these parameters under reasonable ranges (Figures 4.9A-D, 4.10A-D. For detailed implementation see scripts in supplementary files). The simulation of Th prompted us to explore other systems such as sea urchin and frog surfaces, which, while follow roughly similar A, V_C change rules as zebrafish, choose division orientation differently (thus a different Th value), we took advantage of available literature and

general information of these systems (Wray 1997; Chalmers et al., 2003; Wuhr et al, 2008) to validate our model, the example script is provided and the inputs for the three systems are listed in the table below, they produce the corresponding figures. Note that the numbers regarding sea urchin and frog are roughly measured from previous studies and some assumptions may not hold (e.g. sea urchin embryos are known to expand surface during this time (increasing A) and frog embryos have very different cell volumes (animal vs. vegetal)). The inaccuracy of these initial parameters and the rules does not affect the trends of morphogenesis predicted by the model, as these results are only inherently determined by parameters such as Th and γ.

Table S4.2 Initial conditions and parameters used for model simulations.

Step	System	Stage(N)	Th	Sh	γ	Log(T)	$N_S(1)$	V_C (μm^3)	A (μm^2)
1	zebrafish	128	1.25	10	0.8	-4	68	5.00E+05	4.64E+05
2	zebrafish	256	1.25	10	0.8	-4	n/a	2.50E+05	4.23E+05
3	zebrafish	512	1.25	10	0.8	-4	n/a	1.25E+05	4.13E+05
4	zebrafish	1k	1.25	10	0.8	-4	n/a	6.25E+04	4.35E+05
5	zebrafish	2k	1.25	10	0.8	-4	n/a	3.13E+04	4.73E+05
1	frog (*X.laevis*)	32	0.75	10	0.8	-4	32	3.62E+07*	4.50E+06**
2	frog (*X.laevis*)	64	0.75	10	0.8	-4	n/a	1.81E+07	4.50E+06
3	frog (*X.laevis*)	128	0.75	10	0.8	-4	n/a	9.05E+06	4.50E+06
4	frog (*X.laevis*)	256	0.75	10	0.8	-4	n/a	4.53E+06	4.50E+06
5	frog (*X.laevis*)	512	0.75	10	0.8	-4	n/a	2.26E+06	4.50E+06
1	sea urchin	32	0.3	10	0.8	-4	32	4.80E+04***	4.52E+04****
2	sea urchin	64	0.3	10	0.8	-4	n/a	2.40E+04	4.52E+04
3	sea urchin	128	0.3	10	0.8	-4	n/a	1.20E+04	4.52E+04
4	sea urchin	256	0.3	10	0.8	-4	n/a	6.00E+03	4.52E+04
5	sea urchin	512	0.3	10	0.8	-4	n/a	3.00E+03	4.52E+04

* Used cross-section images from Chalmers et al., 2005 to measure and averaged animal and vegetal cells.

** Used radius 600μm to calculate.

**** Used stereoscope images from Wray 1997 to measure assuming radius is 60μm.

Table S4.2 (*Continued*) ***** Used radius 60µm to calculate.

EXTENDED EXPERIMENTAL PROCEDURES

Zebrafish strains and maintenance

The following transgenic strains were used in this study: *Tg(krt4:egfp:caax), tg(actb2:EMTB-3xEGFP), tg(actb2:h2b-tdTomato), tg(actb2:mem-mCherry)* and *tg(actb2:h2b-mCherry2)* have been reported (Wuhr et al., 2010; Krens et al., 2011; Xiong et al., 2013). *Tg(actb2:h2b-egfp/actb2:mem-mCherry2), tg(actb2:h2b-mCherry2/actb2:mem-citrine)* double transgenic lines were generated as described (Xiong et al., 2013). Natural spawning was used and time of fertilization was recorded according to the single cell stage of each clutch. Embryos are incubated in 28ºC during imaging and all other times except room temperature during injections and dechorionating. Staging was recorded using division cycle count and aligned to the normal table (Kimmel et al, 1995). All fish are housed in fully equipped and regularly maintained and inspected aquarium facilities. All fish related procedures were carried out with the approval of Institutional Animal Care and Use Committee (IACUC) at Harvard University.

Microinjections of mRNAs and morpholinos (MO)

1-2 cell stage embryos were injected (Nanoject) 2.3nl 20ng/µl of mem-mCherry2 mRNA with/without 0.1mM cdh1 MO (5-ATCCCACAGTTGTTACACAAGCCAT-3 (Babb and. Marrs, 2004), GeneTools Inc). mRNA is synthesized with the mMESSAGE mMACHINE system (Ambion). Oil drop injection was performed similar using mineral oil.

Timelapse 2-photon/confocal imaging

1-64 cell stage embryos were dechorionated using sharp tweezers (Dumont 55) and mounted animal pole up (or sideways) into a dorsal mount (Megason 2009) with a stereoscope (Leica MZ12.5). Live imaging was performed using a Zeiss 710 confocal microscope (objectives: Plan-Apochromat 20X 1.0 NA, C-Apochromat 40X 1.2 NA) with a home-made heating chamber maintaining 28ºC. Argon laser lines 488nm, 514nm, 561nm and 594nm were used. See also Figure 4.2.

Image Data analysis

Raw data were processed as described (Xiong et al., 2013). Segmentation and tracking were performed using GoFigure2 (www.gofigure2.org) and ACME (Mosaliganti et al., 2012; Xiong et al., 2013). Cell counting was carried out using ImageJ Cell counter plugin after exporting 3D surface rendering images of the raw datasets from FluoRender (Wan et al., 2009). Measurement of L/R ratios was carried out using ZEN (Carl Zeiss) software 3D distance functionality. Measurements were plotted with Microsoft Excel. Movies were generated with GoFigure2, FluoRender2.10 and ImageJ. See also Figure 4.4, Movie S4.2 and corresponding legends for more details.

Modeling

Simple geometry models were drawn on paper and calculated in Microsoft Excel, the full interplay model and its simulations were performed using Matlab (Mathworks). See Figures 4.8,4.10, SI texts.1-6 for details.

Chapter 5

Multibow: Bar-coding Cells for Reliable Lineage Tracing

Fengzhu Xiong, Nikolaus D. Obholzer, Ramil R. Noche, Sean G. Megason

Department of Systems Biology, Harvard Medical School, Boston, 02115, USA

Author Contributions: F.X. and S.G.M. conceived and designed this study. F.X. made the Multibow constructs and generated *elavl3:cerulean-cre* and *actb2:multibow* transgenic lines, performed the experiments and analyzed the data. N.D.O. and R.R.N. generated *hsp70:cerulean-cre* line and provided technical assistance. F.X. and S.G.M. wrote the manuscript.

SUMMARY

We introduce an efficient, noninvasive multicolor labeling method (Multibow) for cell tracing experiments in developmental and regenerative processes. Multibow significantly expands the color code diversity of Brainbow-based approaches and improves the consistency of labeling by using a binary “ON or OFF” Cre-mediated recombination for each fluorescence transgene. We demonstrate the ease of use, high spatial temporal coverage and versatile cell tracking applications of Multibow in zebrafish.

INTRODUCTION

Recent developments in genetics and fluorescent protein technology have allowed elegant designs to label and distinguish cells with multiple colors. In Brainbow (Livet et al., 2007), Cre mediated random recombination on genomic insertions of the Brainbow cassette yields a combinatorial expression profile of different fluorescent proteins (FPs) at different levels. These combinations generate up to ~100 possible visually distinguishable colors for a single cell. This color diversity provides powerful resolution in two main applications: Analyzing detailed organizations of complex structures composed of many cells such as neuronal networks (Livet et al., 2007; Card et al., 2011; Hadjieconomou et al., 2011; Hampel et al., 2011; Pan et al., 2011; Cai et al., 2013), and identifying unambiguously the cells that share clonal origin (Wachsman et al., 2011; Gupta and Poss, 2012; Boulina et al., 2013; Pan et al., 2013). The reliance of both purposes on high resolution imaging renders zebrafish an excellent system of choice. For lineage tracing in zebrafish, Brainbow-like approaches serve as an important alternative method to direct imaging based cell tracking which are still challenging for many tissues (Xiong et al., 2013). Successful adaptations of the original Brainbow design in zebrafish have been reported (Pan et al., 2011; Gupta and Poss, 2012; Pan et al., 2013).

The use of multicolor labeling in cell tracing depends on two key properties of the color generation scheme: First, the diversity of color codes is essential to make an accurate call of clonal vs. non-clonal. The more possible colors a cell may randomly obtain, the less likely its non-clonal neighbors will obtain the same color. Second, the stability of labeling is crucial for identifying the same cells and/or clonal progenies over extended periods of time and/or large

distances of migration. In the original Brainbow design, the color diversity depends on differential signal levels of a few FPs and is therefore limited in total number (Gupta and Poss, 2012; Pan et al., 2013). Moreover, the unique color of a cell/clone is sensitive to factors that affect levels of FP signals, such as promoter activity, which often changes over long time intervals and after changes of cell states (e.g. stress, differentiation, Blake et al., 2006; Mosimann et al., 2011). In addition, acquired signal levels from imaging are often affected by changes in the sample (e.g. depth, autofluorescence). These limitations thus restrict the exploitation of the promising potential of Brainbow-based cell tracing.

RESULTS

A strategy of combining multiple small ON-OFF switch-like brainbow constructs (Multibow)

To expand the labeling diversity and improve stability, we modified the original Brainbow design to a multiple transgene strategy (Multibow). In Multibow, each FP gene adopts a permanent "ON" or "OFF" status upon Cre-mediated recombination (Figure 5.1A, Livet et al., 2007). Multibow provides theoretically 2^{21} possible "digital" color codes (Figure 5.1B) for each single cell by employing 7 FPs of different colors further diversified with 3 different sub-cellular localizations (21 constructs total, Figure 5.1C, Shaner et al., 2007; Ai et al., 2007; Shcherbo et al., 2009), therefore non-clonally related cells at the time of induction are highly unlikely to arrive at the same color code by chance. To test Multibow in zebrafish, we cloned the transgenes to a *beta-actin2(actb2)* promoter construct carrying Tol2 transposon elements (Figure 5.2A), which cause high frequency genomic insertions of *actb2:multibow* upon co-injection of transposase mRNA

(Figure 5.2B, Kawakami et al., 2000). We used heat-shock inducible and tissue specific Cre recombinase driver lines (Figures 5.2C-D, data now shown) to activate recombination in injected embryos and take high resolution images of the same live embryos and larvae at different times (Megason 2009). Multibow injected embryos show bright, mosaic, multiple colors by 16 hours post heat-shock (Figure 5.2E). Signal of nuclear origin has a clear oval shape, while the membrane and cytoplasmic signals can be distinguished using confocal slices through the cell (Figure 5.2F).

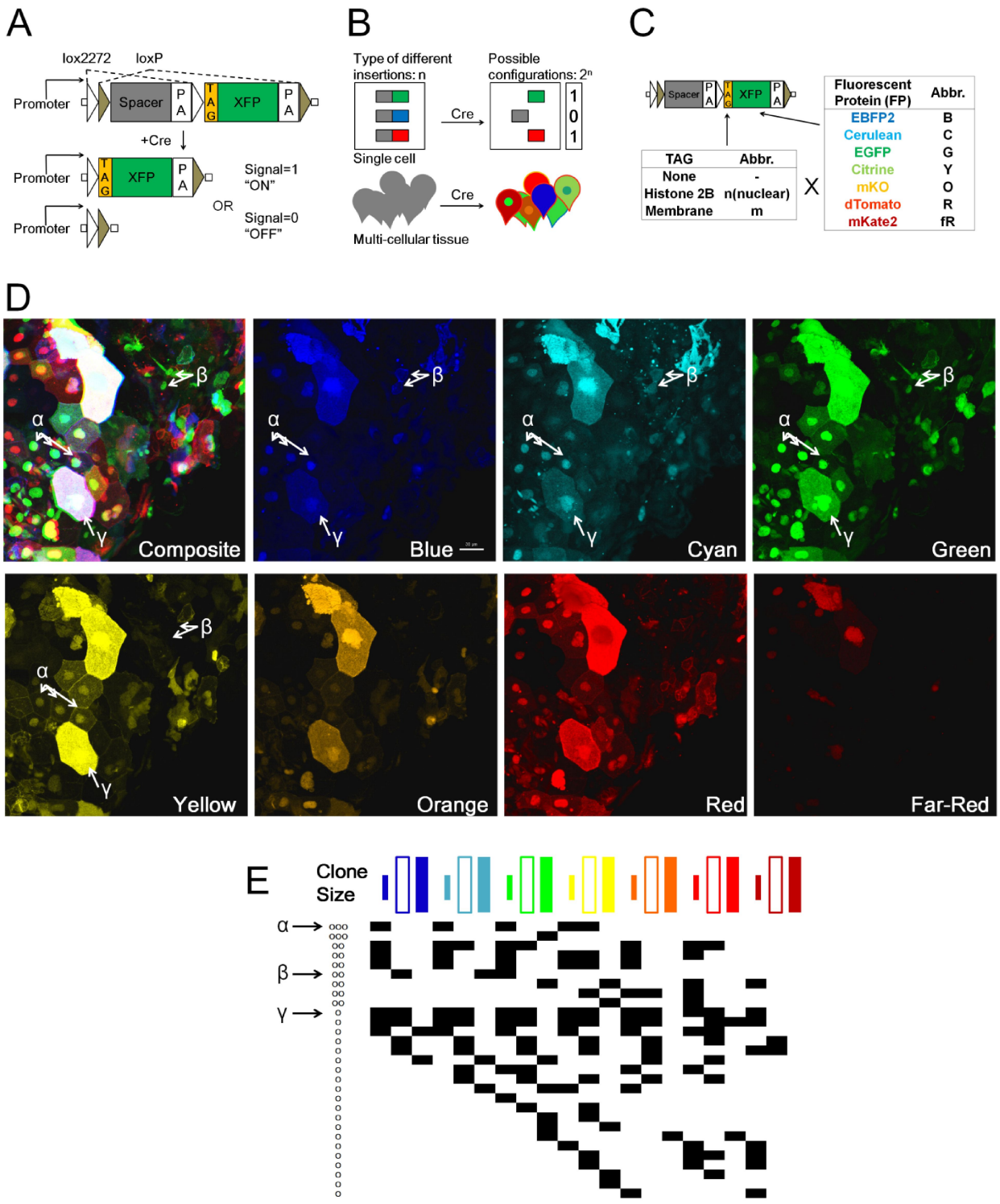


Figure 5.1. Design and test of Multibow in zebrafish

(A) Modified "Brainbow (Livet et al., 2007)" cassette that allows a binary ON/OFF switch.

(B) Multibow Strategy. Each cell harbors multiple different ON/OFF cassettes to generate

Figure 5.1 (*Continued*) random color “digital” barcodes upon Cre-mediated recombination.

(C) Table of Multibow Tags and Fluorescent Proteins (FPs).

(D) Diversity of color codes. Image is a densely labeled region along the trunk of a 40hpf *hsp70:cerulean-cre* embryo injected with all 21 Multibow constructs and heat-shocked at 10hpf. The color and tag diversity generates barcodes for cell clones that appear random and diverse. Intensity differences further help distinguish cells from neighbors visually. The Composite image is made from the green, yellow (turned to blue) and red panels. 3 different clones are highlighted by α, β, γ and corresponding arrows. Scale bar: 10μm. See also Table S5.3.

(E) Partial table of clones of different color codes found in (D). The colored labels indicate nuclear, membrane and cytoplasmic, respectively. A black square indicates this clone being positive for the corresponding color. Distinct "Barcodes" form for different clones. The α, β, γ clones are indicated by arrows.

To assess the extent of color code diversity Multibow can achieve, we examined the emission signal patterns with multiple lasers and filters in embryos injected with all 21 constructs. We found high levels of barcode diversity and randomness (Figures 5.1D-E) as predicted by the design. While it is possible to assess the ON/OFF status for all 7 FPs by different lasers, bleed-through signal from other FPs is unavoidable making certain FPs with highly overlapping emission difficult to separate (e.g. OFP and RFP). In addition, certain FPs (e.g. BFP and CFP) are often masked by autofluorescence of some cells. Therefore, we recommend using 4-channel acquisition (e.g. B/G/Y/R, see Table S5.3 for wavelength ranges), which requires much simpler

and faster imaging set-up but provides sufficient color code diversity (4X3 labels, 4095 barcodes). The full set of 7 FPs thus adds further flexibility to use Multibow with other FP markers (e.g. cell specific reporters).

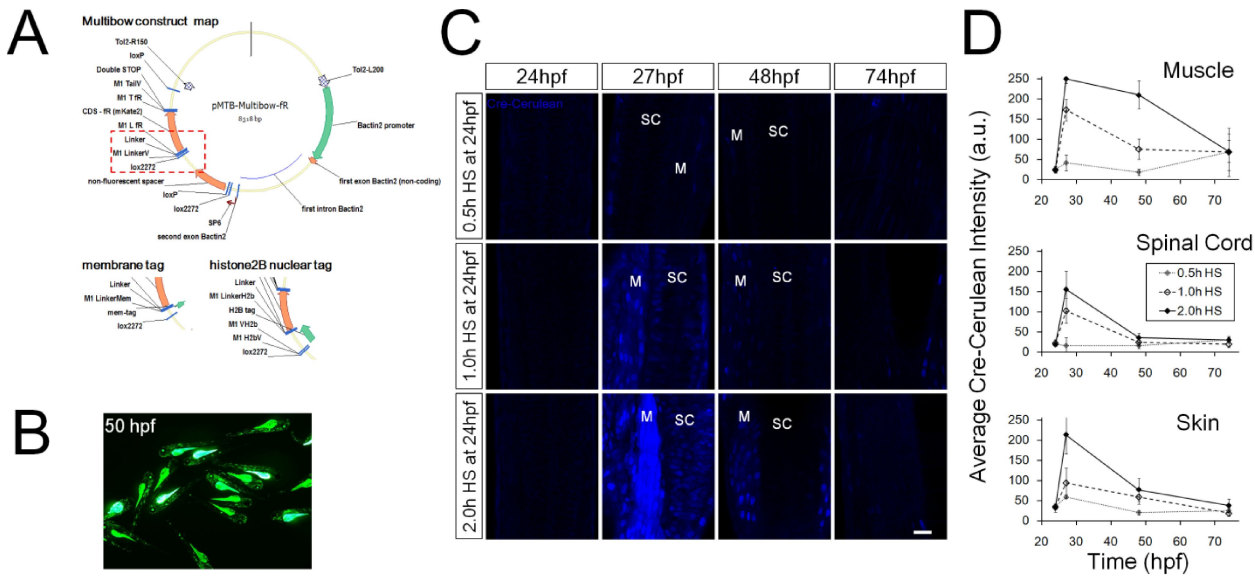


Figure 5.2. Design and Validation of Multibow

(A) Multibow construct map for use in the Zebrafish system. Example construct is pMTB-Multibow-fR(mKate2). Red box: variable region where membrane and nuclear tagged versions are different from the map. See sequence files for details. See also Table S5.1.

(B) Mosaic and uneven distribution of injected DNA. Embryos were injected at 1-cell stage with 20ng/µl pMTB-citrine DNA construct. Yolk is bright with autofluorescence.

Figure 5.2 (*Continued*) (C) Analysis of Cre level dynamics in tg(*hsp70:cerulean-cre*). SC, spinal cord. M, muscle. 0.5h heat-shock does not induce significant Cerulean-Cre expression. Muscle cell expression is more sensitive to heat-shock. Low level expression is also present in some muscle cells without heat-shock. 1.0h heat-shock provides an optimal pulse. Scale Bar: 10μm.

(D) Average fluorescent intensities (+/- s.d.) measured in 10 nuclei from the images in (C).

(E) Onset of Multibow after heat-shock induced Cre expression. The time required for Multibow to become detectable after Cre induction limits its application for lineage tracing in early stage zebrafish embryos before 20 hours post fertilization (hpf). However, the early stages are often more feasible for direct live imaging based lineage analysis (Xiong et al., 2013).

(F) Distinction of membrane and cytoplasmic signals. In 3D projection images (most of the figures) it may be difficult to distinguish signals of membrane and cytoplasmic origin, as bright membrane signal is often detectable in cytoplasm in 3D overlay the signals overlap. Looking through original confocal slices removes most of the difficulties. The top images show a comparison of membrane signal in 3D rendering and slice view, and a slice view of a cytoplasmic signal. The main distinction is a bright edge and fuzzy cytoplasmic signal for membrane FPs and a homogenous signal throughout for untagged FPs. Bottom images show the contrast of neighboring membrane labeled vs untagged FPs of the same color. Difficulties may rise when trying to distinguish the presence of both membrane and cytoplasmic signals of the same FP when the membrane signal is much brighter (See Discussion).

The coverage and stability of Multibow

Spatially, Multibow cells spread over the whole embryo and show excellent diversity in labeled cell types (Figure 5.3A). ~25% of embryos show high cell coverage (>15% of cells labeled by Multibow, Figure 5.4A). Therefore it is easy to find embryos with dense labeling in tissues of interest from the injected pool of embryos (usually >50 per experiment). Temporally, we found persistence of Multibow expression over 10 days (Figure 5.3B) indicating stable genomic insertion of Multibow transgenes, making the strategy feasible to work in older larvae or potentially juvenile and adult animals (Kawakami et al., 2000; Tu and Johnson, 2011).

The design of Multibow suggests that the code should remain the same even if the expression levels of individual FPs change over time. This temporal stability of labeling is crucial to the accuracy of multicolor lineage tracing (Pan et al., 2013). We tested this by following Multibow cells closely over a course of 5 days (Figures 5.3C, 5.4C). The labeled cells indeed show intensity changes of expression between days but the color "codes" determined by the presence of individual tagged FPs are unchanged and the same cells can be identified. In addition, the color codes stay unchanged upon a second heat-shock induction of Cre, suggesting a single 1-hour heat-shock is sufficient to drive ON/OFF recombination to completion (Figure 5.4D). These results validate the design of Multibow in providing stable and discrete color codes for identifying the same cells/clones over time. Together, our data show that Multibow enables easy generation of embryos/larvae that have diversely and stably labeled cell clones suitable for lineage tracing studies.

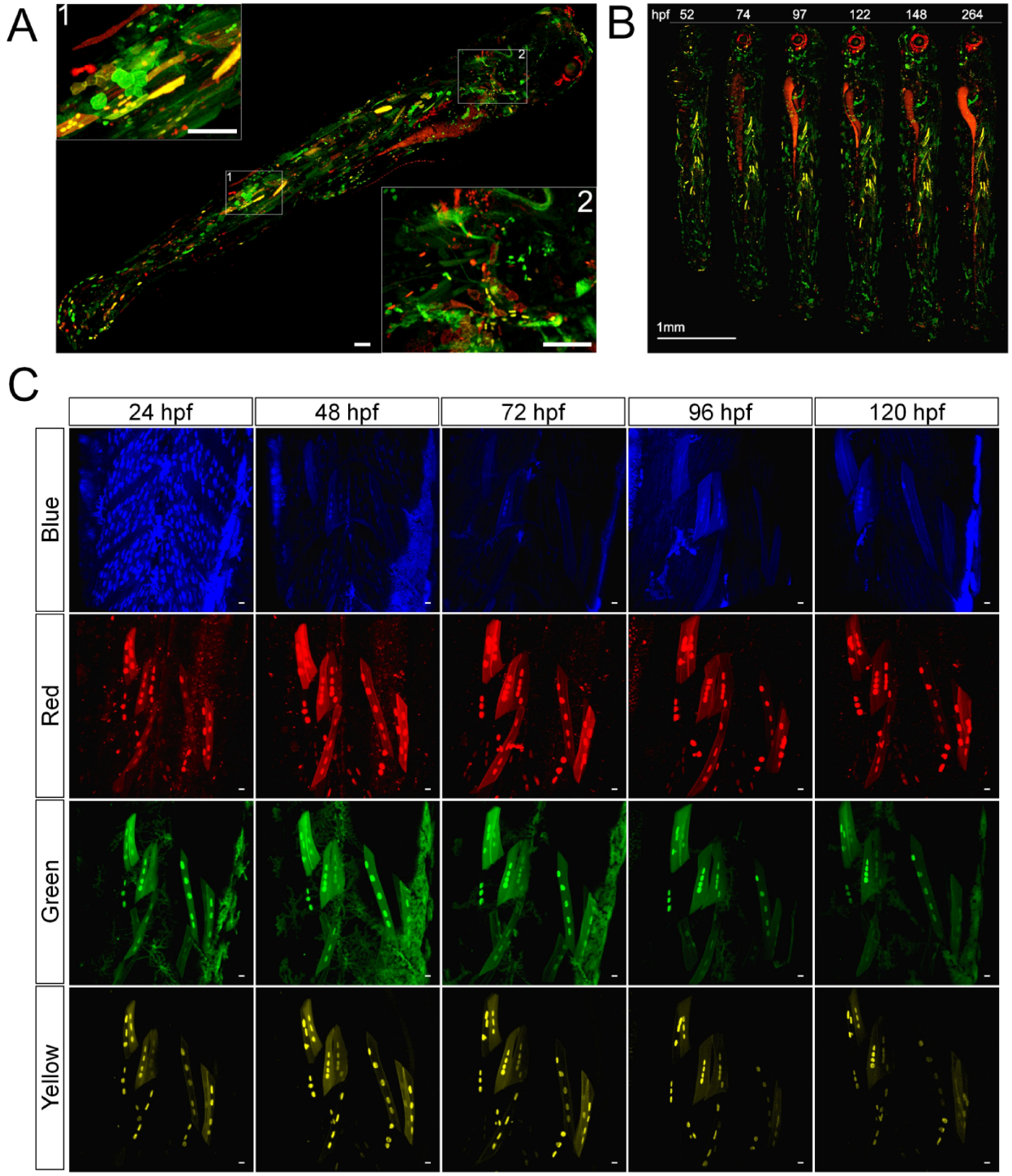


Figure 5.3. Spatial temporal coverage and stability of Multibow labeling

(A) Spatial and cell type coverage of Multibow. A whole 4 day-post-fertilization (dpf) larva was imaged. 2 channels were used for the image with 6 Multibow colors (mR/mG/nR/nG/R/G). Scale

Figure 5.3 (*Continued*) bars: 100µm.

(B) Temporal stability of labeling. The same embryo was imaged once per day to 11dpf. The persistence of labeling indicates genomic insertion of Multibow cassettes. Red patches around the eye and along the gut are auto-fluorescence.

(C) Label stability of color codes to intensity changes. A lateral region of a group of labeled muscle cells were followed for 5 days. The expression profiles of the same cells are invariant during this time despite intensity fluctuations. Four channels were used. Note the 24hpf blue image shows ubiquitous labeling of nuclei by the heat-shock induced Cre-Cerulean expression, which fades away afterwards (See also Figure 5.2C). Scale bars: 10µm.

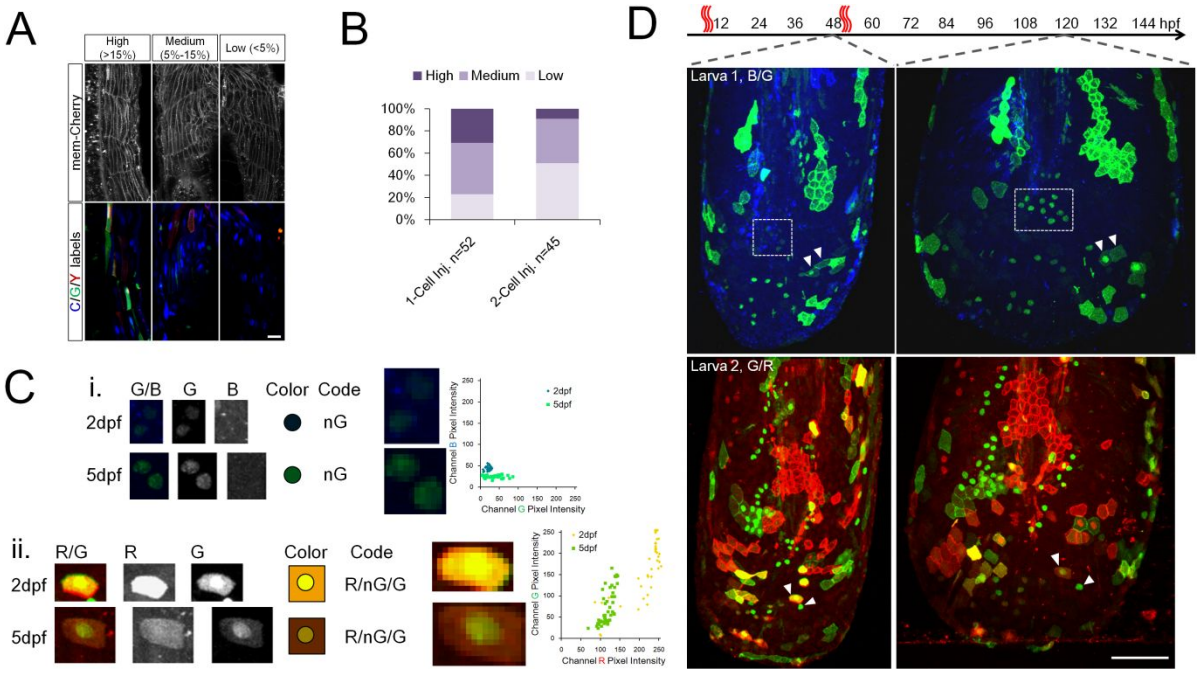


Figure 5.4. Coverage and Stability of color codes

(A) Cell fraction coverage of Multibow analyzed by co-injection with mem-cherry mRNA. The presence of a strong Red ubiquitous marker (mem-cherry) excludes labeled cells by mR/R to be

Figure 5.4 (*Continued*) identified. The actual fraction using all Multibow channels may be higher. Scale Bar: 10μm.

(B) Summary of injected embryos by fraction of Multibow labeled cells. 1-Cell stage injection is critical for higher fraction of labeled cells.

(C) (i) Example cells in which background intensity in Blue channel has dropped while nG intensity increased. (ii) Example cell in which both total and relative intensity of FPs have changed. These cells appear to have changed "color" but can be identified based on the unchanged ON/OFF status of individual tagged FPs. Pixel channel intensity distribution of cells were further analyzed by averaging original images to 1/8 of original pixel number. The RGB intensity values of pixels belonging to the cells were plotted for both time points. The colors of the same cell(s) may change as a result of signal changes between different acquisition times, as reflected by a shift of pixel distribution in the RGB color space.

(D) Heat-shock labeling and imaging were performed as labeled in the timeline. The tissue undergoes growth, and clonal expansion of labeled cells is evident. The color codes remain unchanged (e.g. arrowheads) even though the intensities of expression may change (e.g. Boxes in Larva 1). Some cells labeled at 48hpf have become lost by 120hpf. No new color codes appear after the second heat-shock, indicating completion of recombination by the first heat-shock. Scale Bar: 100μm.

Examples of Multibow applications

We used Multibow to map craniofacial development of a single larva through time by taking one

image stack per day (Figure 5.5A). Many cells and clones are traceable over a 5-day period, an interval that is not currently feasible for performing direct timelapse imaging. The images reveal details of the cell influx that populates the facial structures. We highlight two regions of the "facemap" as examples. In one neuromast (Figure 5.5B), we find that some hair cells share lateral lineage relationships, i.e. progenitors undergo symmetric divisions to give rise to two hair cells. Other hair cells are uniquely coded. These results are consistent with previous work using complex live imaging systems (Swoger et al., 2011). In the inner ear (Figure 5.5C), where extensive morphogenesis movements build the semi-circular canals, we identified initial locations of cells that underwent drastic elongation (>100μm) across the length of the projections. The complex morphological changes of these cells make it challenging to discover them with ubiquitous markers or to trace them with timelapse imaging. In a developing eye (Figure 5.5D), we are able to identify and follow many clones over extended periods of time using a minimal number of acquisitions, even though the tissue has grown considerably and cell locations have changed significantly. We also used Multibow to analyze the cellular origins of regenerated tissue, using the larval tail as an example (Figure 5.5E). Larvae with a suitable number of bar-coded cells in the tail tip were selected and subjected to imaging before, immediately after and 2 days after amputation. The regenerated tissue contains amplified clones of cells of restricted lineages as identified by the same color codes and morphologies, consistent with experiments that use single color mosaic labeling in a large number of animals (Tu and Johnson, 2011).

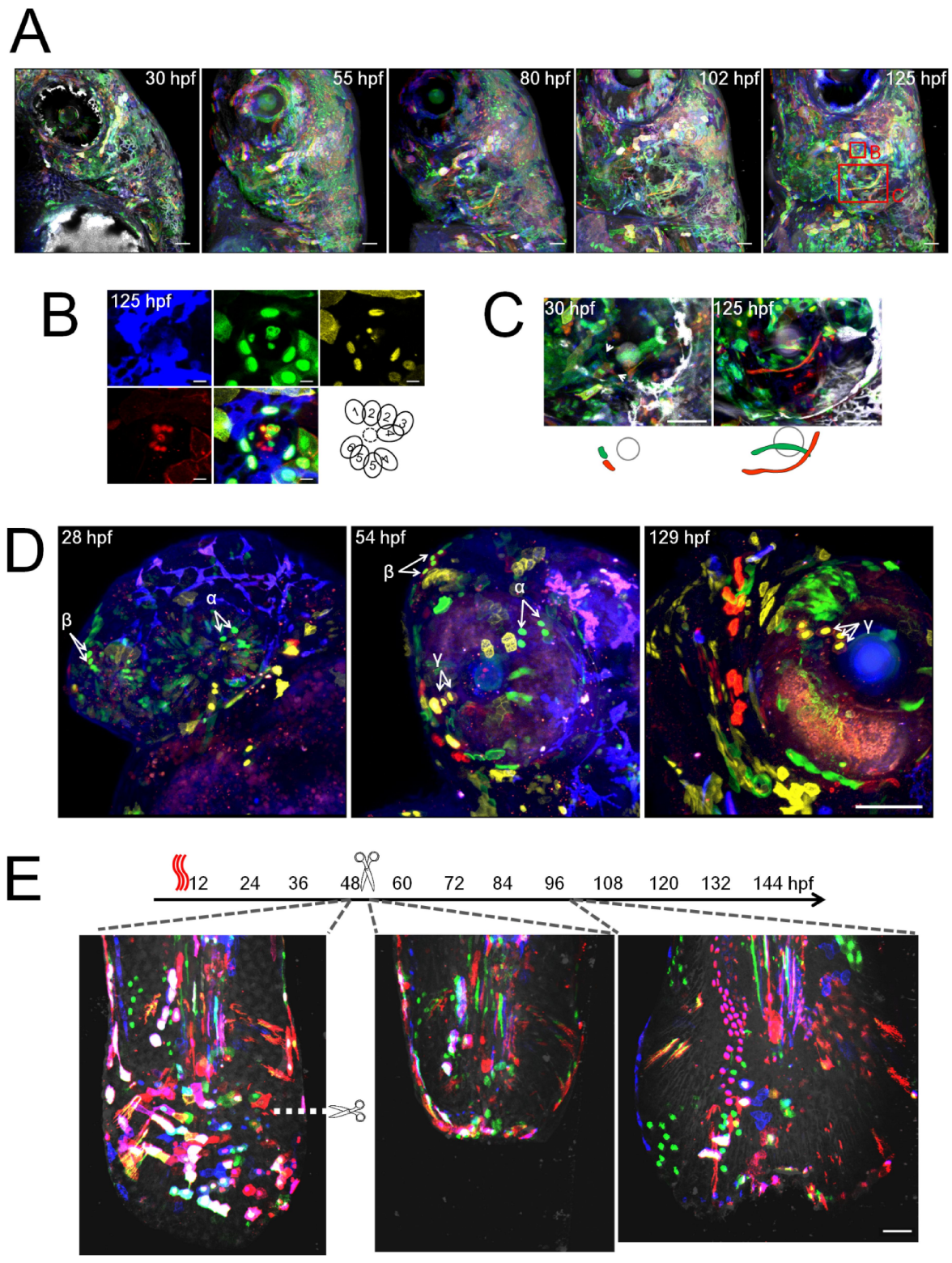


Figure 5.5. Examples of Multibow Cell Tracing in Development and Regeneration

(A) Cranial facial development mapped by Mutlibow. The embryo was heat-shocked at 6hpf. 4

Figure 5.5 (*Continued*) channels (B/G/Y/R) were used. The left face of the larva was imaged. Red boxes: regions highlighted in (B) and (C). Scale bars: 50μm.

(B) Lineage relationship between neuromast hair cells. Dashed line circle indicates the hair bundle. Multibow labeled hair cell color codes: 1(mB/nY/R), 2(mB/mG/nR), 3(mB), 4(nG), 5(mB/nR), 6(R). Scale bars: 10μm.

(C) Identification of cells that undergo remarkable morphological changes during semicircular canal formation. Arrows: initial locations of the two mesenchymal cells that span the projection later. Grey circle: posterior otolith. Scale bars: 50μm.

(D) Clonal expansion near the eye over long time periods. Arrows indicate locations of identified clones α, β, γ. Multiple clones of cells can be seen amplified in number by 129hpf. Scale bar: 100μm.

(E) Multibow analysis of regeneration in the larval tail. Heat-shock labeling, amputation and imaging were performed as labeled in the timeline. Immediately after amputation, the tissue shrank and cells near the wound converged (the images overlay may appear to be slightly out of register due to the changes of the live tissue during the acquisition of different channels, cell identification is not affected as these changes are small and predictable). By 2 days after amputation, most cells that had converged at the frontier of the wound were gone (their unique color codes disappeared). The regenerated tissue came from clonal expansion of cells away from the frontier. These clones show lineage restriction to the original cell type (the morphology of cells in the same clone remains the same). Scale bar: 50μm.

To evaluate the potential of using Multibow in stable transgenic lines, which would have higher percentage of labeled cells, we identified founders harboring single or multiple Multibow transgenes. The recombination functions as expected (Figure 5.6A, data not shown), indicating Multibow transgenes can be used in stable transgenic lines. However, we found while some transgenic lines exhibit random color codes (Figure 5.6B), other lines generate unbalanced and/or reduced diversity in color codes (Figures 5.6C-D). For example, cells expressing one FP become a sub-population of cells expressing another FP, probably due to differential sensitivity and bias of each Multibow locus to the same level of Cre. Although more cells are labeled, stable transgenic lines do not have the first and second levels of randomness discussed above, and as indicated by these data are less powerful for generating color code diversity compared to pooled injections.

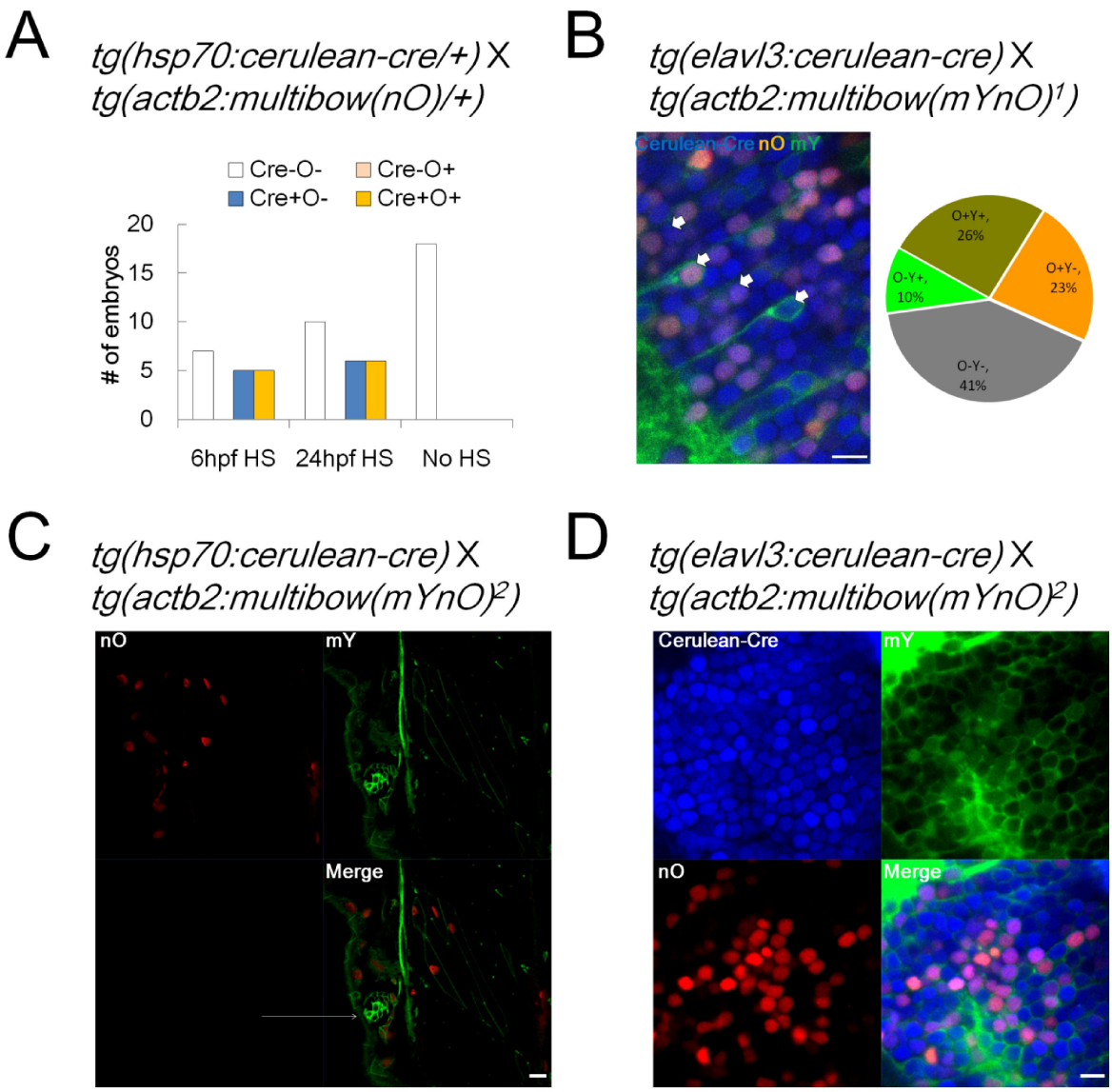


Figure 5.6. Analysis of Multibow transgenic lines

(A) Validation of a single color Multibow line. HS, heat-shock at indicated times (duration: 3 hours).

(B) A double color Multibow line showing ~60% cell coverage and bias towards more nO+ cells. Arrows, example of each of the 4 color codes. Image is a lateral optical slice of the brain, 2dpf. All scale bars: 10µm.

(C) Another double color Multibow line showing mY+ cells cover all nO+ cells, the code of

Figure 5.6 (*Continued*) mY/nO+ is lost in this line. Arrows: mY+/nO- cells. Image is a lateral optical slice of the trunk, 2dpf, heat-shocked at 12hpf.

(D) This line shows very strong coverage of mY+ cells (few mY- cells when Cre expression is constitutive), reducing diversity of color codes. Image is a lateral optical slice of the brain, 2dpf.

DISCUSSION

We have shown Multibow is an easy-to-use, powerful and reliable method for investigating a variety of questions involving cell tracing in the zebrafish system. Multibow brings two improvements to the Brainbow-based multicolor labeling strategy: First, the use of a wider range of FPs and different sub-cellular localization tags allows for the creation of a much larger barcode space, which importantly should all be accessible with similar probabilities by Cre recombination since the constructs are all similar. Second, the use of an ON/OFF indicator makes the color codes more robust to changes in hue, saturation and intensity, which have been the basis for color-codes in previous Brainbow approaches but can be confounded by differential maturation, bleaching, or depth-attenuation of fluorescent labels.

We note that several limitations may restrict the application of Multibow and require consideration when using this strategy. First, the injection of Multibow and the cell divisions that follow cause uneven inheritance of the injected pool, which makes the labeling mosaic. While mosacisim adds an advantage for lineage tracing as more separated clones are easier to identify and follow, it reduces the throughput of experiments. Less information can be recovered than densely labeled embryos and many embryos may need to be imaged and analyzed to obtain

statistical lineage information. Second, a delay of labeling onset exists between the induction of Cre activity and detectable fluorescence, during which the cell may divide, so it is not practical to be exact on when the cell/clone obtains its barcode. Improvement on the Cre activity control (e.g. unstable Cre, drug inducible activity, etc) may enhance the temporal accuracy of labeling. Third, in certain cell types or situations the sub-cellular localization tags may become difficult to separate. For example, in neurons that are densely packed the cytoplasmic signal may often be covered by a strong membrane signal. In cells with a bright cytoplasmic signal and weak membrane signal of the same FP, the membrane signal may not be distinguished during analysis. If this relative brightness changes over time, a barcoding error may appear. Therefore the power of Multibow is reduced in these cases and more careful analysis is needed.

Despite the discussed limitations, we believe Multibow is scalable for high-throughput mapping of development and regeneration by using automated image analysis approaches (Mosaliganti et al., 2012), and is also adaptive for use with different promoters, Cre drivers and transgenic markers that are available to study specific problems. Multibow is also potentially applicable in other model organisms where high genomic insertion rate is technically achievable (Urasaki et al., 2008; Casco-Robles et al., 2010).

EXPERIMENTAL PROCEDURES

Multibow cloning and preparation.

To clone Multibow constructs, the original pCMV-Brainbow-1.1 (Livet et al., 2007) construct was digested with NheI/XmnI and the 3.4kb fragment was cloned into a pMTB vector containing

the Tol2 elements and the beta-actin2 (*actb2*) mini gene driver (Figure 5.2A, Xiong et al., 2013) at the SpeI/ZraI sites. The resulting construct was digested with BstBI/BbsI and re-ligated to truncate both the original mem-cherry and Kusabira orange sequences to generate the non-fluorescent spacer sequence flanked by loxP and lox2272 sites. The resulting construct served as backbone for cloning of different FP sequences at the original mem-EYFP site. To facilitate modularity and efficiency of cloning, small linker sequences were used to connect the backbone to different tags/FPs and between tags and FPs. Primers joining linker sequences and FP/backbone sequences (See Table S5.1 for the list of FPs used and other sequences) were used to amplify the vector, FPs and tag fragments using a Fusion High-Fidelity PCR Mix (NEB). Fragments were combined in an isothermal assembly reaction (Gibson et al., 2009, Table S5.2) to generate Multibow constructs. Reaction mixes were transformed with 5-alpha F'Iq cells (NEB) and screened by colony PCR. Multibow constructs displayed instability in *E.coli* cultures and had low yield in Maxi/Midi-preps. Mini-prep was used to harvest Multibow constructs. A further purification step using the MinElute PCR purification kit (Qiagen) was applied to clean and concentrate Multibow constructs. Using isothermal assembly, the Multibow cassettes can be efficiently cloned into other vector backbones. To generate the *elavl3:cerulean-cre* construct for transgenic lines, the elavl3(HuC) promoter (Park et al., 2000) was cloned to replace the actb2 promoter in an *actb2:cerulean-cre* construct. To generate the *hsp70:cerulean-cre* construct for transgenic lines, the *hsp70* promoter was cloned to replace the actb2 promoter in a *actb2:cerulean-cre* construct. The Cerulean-Cre fusion protein allows real time monitoring of presence of Cre in live embryos.

Zebrafish maintenance, injection and transgenic lines.

All fish are housed in fully equipped and regularly maintained and inspected aquarium facilities. All related procedures were carried out with the approval of the Institutional Animal Care and Use Committee (IACUC) at Harvard University. AB wild-type and RNF pigmentation mutant strains were used. The transgenic strains (tg(*elavl3:cerulean-cre*), tg(*hsp70:cerulean-cre*) and tg(*actb2:multibow*) lines) were generated by standard injection/screening methods. Natural spawning was used to obtain embryos and time of fertilization was recorded according to the single cell stage of each clutch. Embryos were incubated in 28ºC during imaging and all other times except room temperature during injections and dechorionating. To make the transposase mRNA, a pCS-TP (Kawakami et al., 2000) construct was linearized with NotI and purified with the MinElute PCR purification kit (Qiagen). The construct was then used to synthesize transposase mRNA with the mMESSAGE mMACHINE system (Ambion). The injection solution contained 30ng/µl transposase mRNA and 20ng/µl mixture of Multibow constructs (equal quantity of each single construct). The combinations of Multibow constructs were chosen depending on experimental needs. 1-cell stage (1-cell stage is important for good labeling coverage) embryos were injected (Nanoject) with 2.3nl injection solution and were screened (MVX fluorescent macro scope, Olympus) for health before applying heat shock (37ºC air incubator, timing and duration depend on experimental needs). The injected embryos show slight delay of development as expected from injection experiments. No significant health difference was observed between Multibow injected embryos and control embryos. Amputation of the tail was performed using a clean razor blade on anesthetized (0.2% Tricaine) embryo/larva on

agarose mounts.

Imaging and image analysis.

Embryos showing Multibow expression were screened and selected with MVX fluorescent macro scopes (Olympus) and then mounted to imaging molds as described (Megason 2009). Z-stacks were taken using a Zeiss 710 confocal microscope and Zen software with a home-made heating chamber maintaining 28ºC. Different laser and filter settings were used for different FPs (Table S5.3). Bleed through signal is common with FPs that have overlapping emission spectra, but can be distinguished based on signal pattern in adjacent channels using a different excitation laser and/or filter. Barcodes were assigned based on the presence of signal in individual channels. Single channel original images were analyzed in Zen (Zeiss) through Z-slices and rendered into 3D composite views with FluoRender 2.10 (*http://www.sci.utah.edu/software/127-fluorender.html*). The 3D images were further overlaid in Photoshop to make composite multicolor figures.

ACKNOWLEDGMENTS

The authors thank J. Sanes for information on the Brainbow constructs, H. Higashijima for *elavl3* (HuC) construct, A. Pan, I. Swinburne, J. Apfeld, A. Klein and C. Cepko for discussion. This work was supported by R01 DC010791 from NIH. F.X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

CHAPTER 5 REFERENCES

Ai, H.W., Shaner, N.C., Cheng, Z., Tsien, R.Y., and Campbell, R.E. (2007). Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins. Biochemistry *46*, 5904-5910.

Blake, W.J., Balazsi, G., Kohanski, M.A., Isaacs, F.J., Murphy, K.F., Kuang, Y., Cantor, C.R., Walt, D.R., and Collins, J.J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. Mol. Cell *24*, 853-865.

Boulina, M., Samarajeewa, H., Baker, J.D., Kim, M.D., and Chiba, A. (2013). Live imaging of multicolor-labeled cells in Drosophila. Development *140*, 1605-1613.

Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W., and Sanes, J.R. (2013). Improved tools for the Brainbow toolbox. Nat Meth *10*, 540-547.

Card, J.P., Kobiler, O., McCambridge, J., Ebdlahad, S., Shan, Z., Raizada, M.K., Sved, A.F., and Enquist, L.W. (2011). Microdissection of neural networks by conditional reporter expression from a Brainbow herpesvirus. Proc. Natl. Acad. Sci. U. S. A. *108*, 3377-3382.

Casco-Robles, M.M., Yamada, S., Miura, T., and Chiba, C. (2010). Simple and efficient transgenesis with I-SceI meganuclease in the newt, Cynops pyrrhogaster. Dev. Dyn. *239*, 3275-3284.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A.,3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343-345.

Gupta, V., and Poss, K.D. (2012). Clonally dominant cardiomyocytes direct heart morphogenesis. Nature *484*, 479-484.

Hadjieconomou, D., Rotkopf, S., Alexandre, C., Bell, D.M., Dickson, B.J., and Salecker, I. (2011). Flybow: genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. Nat. Methods *8,* 260-266.

Hampel, S., Chung, P., McKellar, C.E., Hall, D., Looger, L.L., and Simpson, J.H. (2011). Drosophila Brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. Nat. Methods *8,* 253-259.

Kawakami, K., Shima, A., and Kawakami, N. (2000). Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. Proc. Natl. Acad. Sci. U. S. A. *97,* 11403-11408.

Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450,* 56-62.

Megason, S.G. (2009). In toto imaging of embryogenesis with confocal time-lapse microscopy. Methods Mol. Biol. *546,* 317-332.

Mosaliganti, K.R., Noche, R.R., Xiong, F., Swinburne, I.A., and Megason, S.G. (2012). ACME: automated cell morphology extractor for comprehensive reconstruction of cell membranes. PLoS Comput. Biol. *8,* e1002780.

Mosimann, C., Kaufman, C.K., Li, P., Pugach, E.K., Tamplin, O.J., and Zon, L.I. (2011). Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. Development *138,* 169-177.

Pan, Y.A., Freundlich, T., Weissman, T.A., Schoppik, D., Wang, X.C., Zimmerman, S., Ciruna, B., Sanes, J.R., Lichtman, J.W., and Schier, A.F. (2013). Zebrabow: multispectral cell labeling for cell tracing and lineage analysis in zebrafish. Development *140,* 2835-2846.

Pan, Y.A., Livet, J., Sanes, J.R., Lichtman, J.W., and Schier, A.F. (2011). Multicolor Brainbow imaging in zebrafish. Cold Spring Harb Protoc. *2011,* pdb.prot5546.

Park, H.C., Kim, C.H., Bae, Y.K., Yeo, S.Y., Kim, S.H., Hong, S.K., Shin, J., Yoo, K.W., Hibi, M., Hirano, T., *et al.* (2000). Analysis of upstream elements in the HuC promoter leads to the establishment of transgenic zebrafish with fluorescent neurons. Dev. Biol. *227,* 279-293.

Shaner, N.C., Patterson, G.H., and Davidson, M.W. (2007). Advances in fluorescent protein technology. J. Cell. Sci. *120,* 4247-4260.

Shcherbo, D., Murphy, C.S., Ermakova, G.V., Solovieva, E.A., Chepurnykh, T.V., Shcheglov, A.S., Verkhusha, V.V., Pletnev, V.Z., Hazelwood, K.L., Roche, P.M., *et al.* (2009). Far-red fluorescent tags for protein imaging in living tissues. Biochem. J. *418,* 567-574.

Swoger, J., Muzzopappa, M., Lopez-Schier, H., and Sharpe, J. (2011). 4D retrospective lineage tracing using SPIM for zebrafish organogenesis studies. J. Biophotonics *4,* 122-134.

Tu, S., and Johnson, S.L. (2011). Fate restriction in the growing and regenerating zebrafish fin. Dev. Cell. *20,* 725-732.

Urasaki, A., Mito, T., Noji, S., Ueda, R., and Kawakami, K. (2008). Transposition of the vertebrate Tol2 transposable element in Drosophila melanogaster. Gene *425,* 64-68.

Wachsman, G., Heidstra, R., and Scheres, B. (2011). Distinct cell-autonomous functions of RETINOBLASTOMA-RELATED in Arabidopsis stem cells revealed by the Brother of Brainbow clonal analysis system. Plant Cell *23*, 2581-2591.

Xiong, F., Tentner, A.R., Huang, P., Gelas, A., Mosaliganti, K.R., Souhait, L., Rannou, N., Swinburne, I.A., Obholzer, N.D., Cowgill, P.D., Schier, A.F., and Megason, S.G. (2013). Specified neural progenitors sort to form sharp domains after noisy shh signaling. Cell *153*, 550-561.

Supplementary Tables (Legends only, see attached supplemental files)

Table S5.1. Sequences Table.

Includes primer sets, tag sequences and FP sequences used to assemble Multibow constructs. See also sequence files, Figure 5.2A.

Table S5.2. Modified Isothermal Assembly Ingredients Table.

Mix fragments at equal molecular ratio in 5µl and mix with 1 Isothermal Assembly reaction aliquot (15 µl), then incubate at 50ºC for 15-30 minutes. Use 1µl reaction mix to transform competent cells. See also Gibson et al., 2009.

Table S5.3. Confocal Imaging Settings Table.

Applying additional laser lines at different wavelengths will help distinguish the presence of different FP genes. See also Figures 5.1D-E.

Chapter 6

Ideas and Perspectives

Fengzhu Xiong

Department of Systems Biology, Harvard Medical School, Boston, 02115, USA

Author Contributions: F.X. wrote the manuscript.

This dissertation provides groundbreaking work of cell tracking analysis in complex developmental processes, demonstrating that the improvement in observation not only can directly lead to more definitive and quantitative descriptions, but also allows testing of different models of conceptual significance. This is likely a fruitful direction in the near future of developmental biology for several reasons: First, recent years have seen a quick advancement of microscope designs that enhance resolution, coverage and throughput at the same time (Keller 2013). These properties, which have been pursued individually by high level of specialization in the past, are all important for conducting live observations of development in a systematic manner (Megason and Fraser, 2007). Second, the knowledge of genomes, gene editing tools and fluorescent proteins have arrived at a level of sophistication that allows quick and clean generation of specific live labels in cells and embryos (Giepmans et al., 2006; Hwang et al., 2013;

Cong et al., 2013; Zu et al., 2013; Howe et al., 2013). This will increase the range of problems that could be addressed by direct observation. Third, image data storage and analysis are currently topics of focus in the fields of computer sciences and software engineering. Accompanied by the fast hardware development of computational and sharing capabilities, processing of large data sets such as those we generated in this work (Chapters 2-4) will likely become more efficient, automated and standardized. Fourth, systems biology has in recent years laid solid foundations for constructively understanding protein/gene networks and cell behaviors (Alon 2007). Multicellular development adds a new layer of complexity and is likely to be the next target of novel theoretic work (Chapter 4 serves as a simple example along this line of thinking). Together these trends make a convincing argument for further developing and applying the methodology and philosophy of the current dissertation in future work. Here I discuss the perspectives and general ideas of cell tracking and developmental paradigms that conceptually summarize/extend the results of this work.

Principles of cell tracking in studying development

Cell tracking is among the most powerful tools to provide insights in the study of developmental biology. As demonstrated in chapter 2, registering cellular events on a common time axis allows direct tests between "position → fate" vs. "fate → position" models. The reliability of cell tracking is essential for drawing solid conclusions. However, in practice the "Trackability" of a certain experimental system may just fall near the borderline where whether the same cell was tracked was difficult to determine. To improve the trackability for those systems and to evaluate or

design a system's trackability in general, it requires a good understanding of the principles of cell tracking. To provide a common theme for all chapters of this dissertation, here I propose a conceptual analysis of trackability. The action of tracking is to recognize the same object after a certain amount of time. The validity of any tracking is checked in two ways: First, whether the object identified at a later time retains an observable property that is consistent with (Consistency of Labeling) and specific to (Density of Labeling) the object identified at an earlier time. For example, a name, a fingerprint and a DNA profile of a person can all be the observable properties that allow tracking of the person in a 1-million people population over a 10-year time interval. However, the Consistency of Labeling of these properties varies. The person may for some reason change his/her name during the 10 years. A less likely change could happen with the fingerprint such as injury to the fingers. The DNA profile, compared to the other two, is the most consistent property therefore offers the highest trackability. The other important factor is the Density of Labeling, which is defined as how often the observed property is the same in other individuals in the population where the individual of interest is being tracked. The same name may reappear many times in a 1-million population therefore this property has a high Density of Labeling. The DNA profile has a much lower Density of Labeling value, however, it may cause a wrong track if the person being tracked has a monozygotic twin. In this regard, the fingerprint is the most specific observable property. Combining the Consistency and Density of Labeling we obtain a measure of the individuality of an observable property that correlates with the reliability of tracking. Second, whether the observation is frequent enough to establish a continuum of records of the object being tracked, which means that the difference between two consecutive

observations is so small that no other scenario is possible except the object identified being the same. An object may change all of its observable properties (low Consistency of Labeling) but a frequent enough sampling still validates the tracking. For example, a video camera set-up near a caterpillar undergoing metamorphosis will validate the tracking of the individual later as a butterfly. Frequent observation also allows tracking at high Density of Labeling. For example, high speed cameras recording a migrating swarm of bees allow the analysis of individual bee trajectories even though the label used in this tracking process (a bee image) is found in all individuals of a densely packed population. In practice, a combination of different properties are observed to increase trackability and the frequency of observation is often optimized to be as high as possible under technical limitations. Cell tracking in development serves as a good example where these concepts aid the choice of system of interest and experimental design.

Here I define the above-mentioned tracking-related properties of a developmental system as: the Frequency of Observation ($\boldsymbol{f}$), the Consistency of Labeling ($\boldsymbol{C_i}$) and the Density of Labeling ($\boldsymbol{D_i}$), in which $\boldsymbol{C_i}$ and $\boldsymbol{D_i}$ compile the Individuality ($\boldsymbol{I}$) of the cell ($\boldsymbol{I = \Sigma\, C_i / D_i}$) and the Trackability ($\boldsymbol{T}$) of the cell is defined as $\boldsymbol{T = I \cdot f}$. Here $\boldsymbol{C_i}$ represents the stability of one observable property of a cell and there are i=1,2,3...n such properties (Same with $\boldsymbol{D_i}$). With this formalization it becomes clear how cell tracking methods used in this dissertation and other studies essentially work: First, direct high resolution imaging increases $\boldsymbol{f}$ (for example, chapter 2). Second, labeling cells with fluorescent proteins provides an observable property that has a high $\boldsymbol{C_i}$. Finally, using mosaic and multiple fluorescent labeling reduces the $\boldsymbol{D_i}$ of the fluorescent labels (for example, chapters 3 and 5). These principles show us the considerations when

designing a cell tracking method and interpreting experiments that have an inherent cell tracking component. For example, one may assume the position of cells does not change and use it as a high $\boldsymbol{C_i}$ to make arguments on other observed changes, such as gene expression. The stable position assumption here must be validated by introducing a more reliable label, for example, introduced mosaic fluorescent labels. In practice, obtaining a consistent property for cells is not difficult. The more challenging aspects are increasing $\boldsymbol{f}$, and reducing $\boldsymbol{D_i}$ while keeping a good throughput. Microscopy technology continues to push the frontier of temporal resolution, and labeling strategies such as Brainbow and Multibow (chapter 5) generate more unique labels. With these developments it is conceivable that cell tracking will become more broadly applied to study developmental problems in the near future.

Patterns and forms as emerging properties of cell-cell, cell-environment interactions

Development is a collective process of many individual cells. This adds a key layer of complexity to the biochemical nature of cell behaviors. From the tradition of biochemistry and genetics, much emphasis has been placed on the genome encoding the blueprint of development. However, a genome does not generate an embryo outside the context of a functional egg and the subsequent interactions between cells that multiplied from it, or the environment the embryo develops in. An analogy to this is the development of a social structure *de novo* in an island-survivor situation. A group of people with no specialized training of any kind (naive progenitor cells) arrive at an isolated island (confined environment such as an egg). Initially, individuals respond to stimuli and interact with the environment following a similar set of survival rules

(looking for food, finding shelters, etc), just the same as cells predictably respond to a defined cue by using the same genome. Over time, however, differences and organizations start to emerge (specification and patterning). An individual with slightly higher level of curiosity (intrinsic) or simply happening to be in a specific location of the environment (extrinsic) may find and maintain a habitable cave on higher grounds of the island, whereas an individual who first catches a fish near the water may become the leader of the fishermen. Trades will then emerge between fishermen and cave occupiers such as exchanging fish for shelter thus functional social structures form.

This example highlights several interesting properties of development: First, the response system (genome) of individuals (cells) to others and the environment is important. Perturbing this system may lead to failure to form a viable social structure (disabling arms of all individuals), or the formation of a mutated one (removing language), or a normal one (shaving hairs). Note that these perturbations alter eventual developmental outcome via interactions that depend on the logic of the rest of the response system and the environment. In this context the picture that a gene "instructs/determines" a certain outcome is misleading. Development is changed because the perturbation altered a component that participates in the interactions, identifying the component does not answer what the nature of the interactions is. Only by observing the perturbed processes can we learn how the component fits in the picture. Second, the initial conditions are essential yet the developmental outcome is robust to reasonable amounts of variations in these conditions. A fishing and cave-dwelling pattern would likely emerge as long as there are enough individuals to divide labor, accessible caves and fish. The

developmental process itself may be quite variable under different conditions and it may not be the same individuals who become fishermen if the experiment is repeated with the same initial conditions (variable lineages). This suggests that the interactions that drive development have been evolutionarily selected to be inherently "canalizing" (Waddington 1942), setting a fundamental constraint on the ways by which the interactions can be organized. In chapter 4 we have shown that the feedback between cell shape and division orientation would underlie a developmental process that makes the right cell shapes. The existence of this feedback is unlikely just by chance but instead a necessary way to canalize variable initial conditions. Other mechanisms, such as tissue expansion or cell growth, may be able to produce similar outcomes but fall short in providing robustness. Third, the strategy of development by interactions along the way allows the process to pursue optimality, which is a selectable property. To maximize the efficiency of the society, the number of fishermen vs. cave-keepers will depend on the demand and supply of fish and caves and is not optimal to plan ahead or pre-determine. A notable example of developmental optimality is the phenomenon of "scaling" in which embryos/tissues arrive at correctly proportioned patterns but variable sizes after perturbations. For example, in zebrafish, if a portion of the embryonic yolk is removed early, the embryo will later become smaller compared to unperturbed siblings at the same stage. Again, this optimality is achieved by complex interactions that allow the cells to measure the changed environment and remake group decisions. Under this principle, it is conceivable that the genome has evolved logics that produce cell responses that "anticipate" changes in environment in order to adjust to optimality if they do occur.

These concepts are useful in future studies of development. For example, perturbations (e.g. mutations) that do not produce an apparent or dramatic phenotype, or that only produce partially penetrant phenotypes are more interesting in the light of the robustness principle (instead of being considered as not developmentally relevant or not important). By observing closely at the ones that are penetrant (develop the phenotypes) vs. the ones that are not, it is likely differences of cell behaviors, gene expression levels and other properties of the system will be identified thus revealing the key interactions that decide the different outcomes in a population. It is also plausible to evaluate available models on their robustness and optimality and hypothesize the existence of additional mechanisms that provide these properties if current models do not. Using these ideas when analyzing a developmental problem is likely to inspire guided discoveries and tests.

The unavoidability and usefulness of noise in development

Noise arises from random fluctuations of molecules/structures at small numbers that is physically impossible to control (Paulsson 2004). In addition to this basic type of noise, biological processes are confined by a lot of restrictions in design which bring different kinds of noises (Eldar and Elowitz, 2010). For example, most enzymes that organisms use are sensitive to temperature which is a locally fluctuating property, generating noise in enzymatic activities. The ubiquitous presence of noise constrains the design of biological systems further but can also serve as a way of understanding mechanisms of development. As the biological processes move to higher levels, e.g. from biochemical interactions between a few proteins to whole networks

and then to multicellular populations, the number of variables that affect the process increases exponentially. To avoid butterfly effect and make developmental processes robust to a wide range of variables, it is intuitive to evolve regulations of noise at different levels. In chapters 2 and 3 we have demonstrated how the mechanism of cell sorting can ensure pattern precision in the face of positional and signaling noises that are difficult to avoid in a dynamic tissue.

It is less intuitive, however, to consider the usefulness of noise in development. One possibility of the role of noise is to produce diversity that initiates/participates the breaking of homogeneity in the population. This has been implied to play a role in developmental patterning since Turing (Turing, 1952). One interesting example relating to this role of noise comes from the fate specification in the ventral neural tube. It was found that, intriguingly, in double knockout mice of Shh and Gli3 (a repressor of neural progenitor specification), the ventral fates become specified in disorganized and mixed pattern (Wijgerde et al., 2002). This shows that Gli3 is in place to repress noise in the expression of Shh-downstream fate regulators, which are known to initiate their own positive feedback loops to establish stable fates that do not depend on further Shh signaling (Lek et al., 2010). The repressing mechanism by Gli3 reduces the sensitivity of the cells to Shh and seems to be the key that allows both Shh to specify the correct pattern and the specified cells to robustly maintain their fates. This example shows both noise regulation and the role of noise to initiate a cascade of events. It implies that there is a basic level of noise in the Shh interpretation gene regulatory network (GRN), which we observed in chapter 3. Another possibility of the usefulness of noise is that noise can produce a range of states in the population, which may be important for developmental mechanisms to act on. This is best demonstrated in

the distribution of cell shapes we found in chapter 4. The broad distribution produces Surface-Deep divisions in the population at each cell cycle that gradually flattens the average cell shape. If this mechanical noise does not exist, the cell shapes become homogeneous and will divide in the same orientation following the division rule. In this scenario the divisions may become oscillatory between two orientations from cycle to cycle, which is unlikely to be robust or optimal. Together, these example argue that the characterization of noise is potentially revealing of the developmental strategies used in the system. Noise, once confirmed to be biological instead of experimental, should be assessed carefully for its role and not disregarded as "imperfections" of a presumed deterministic mechanism. Our understanding of noise in development will become better as the measurement methods become higher resolution and more quantitative.

Achieving global behaviors by local regulations

A general theme of development is to produce patterns and forms over large length scales that far exceed the size of a cell. While some cells may influence others over several cell diameters (e.g. morphogen gradients), the difference in scale is still large enough that it is reasonable to assume that cells are information restricted to its neighborhood and "know" little about what most other cells in the population are doing. This sets a challenge for the design of developmental mechanisms: how to orchestrate a specific group behavior given a lot of similar individuals that do not see further than their neighbors?

Environmental conditions set-up by the mother provide some answers. For example the

morphogens deposited in the egg could form a far-reaching gradient due to their high levels of concentration compared to the lower levels produced later in development by small groups of cells. In addition, it is possible that the logics of simple, local rules that can evolve or be regulated by the cells collectively determine the behavior of the population as cell interact with each other and with the environmental constraints (chapter 4). I suggest that more complex types of morphogenesis or patterning processes may also be based on similarly simple local rules that apply to all the cells. This implies a strict constraint on the types of mechanisms that could emerge. One notable example is the finding that cell divisions are often essential for morphogenesis. Oriented divisions take advantage of a basic proliferation machinery and creates local patterns by controlling the locations and shapes of daughter cells. This process eventually may generate enough force to deform the whole tissue to make novel global shapes. Oscillations of gene expression that can be coupled with neighbors is another way to coordinate cells over large length scale. A cell with coupled oscillation receives information from the far away source in the form of the cell's phase. Here a GRN oscillation and a delay that are both fixed properties for all cells make a clock that relays information.

Together, this idea suggests that the evolution of development may have exploited all possible single cell behaviors to produce novel structures and patterns. The examples discussed here are the successful ones that development still heavily relies on. Future studies of poorly understood developmental and regenerative processes may benefit from the thinking that the population of cells is a society of simple and similar individuals that have limited communicative ability. This principle is likely also true when considering biological systems at other levels, for example the

coordination of molecular interactions to produce cellular behavior, or the inter-species interactions that shape the eco-system. The generation of global behavior from local interactions is a realm of developmental biology where systematic observations and theoretic models are likely to continue to be powerful and productive.

ACKNOWLEDGMENTS

The author thanks S. Megason for advisory, C. Cepko, J. Flanagan, M. Levin, A. Lassar and B. Olsen for helpful discussions. This work was supported by R01 DC010791 from NIH. F.X. is also supported by the graduate program of Biological Sciences in Dental Medicine at Harvard University.

CHAPTER 6 REFERENCES

Alon, U. (2007). Network motifs: theory and experimental approaches. Nat. Rev. Genet. *8*, 450-461.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. Nature *467*, 167-173.

Giepmans, B.N., Adams, S.R., Ellisman, M.H., and Tsien, R.Y. (2006). The fluorescent toolbox for assessing protein location and function. Science *312*, 217-224.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. Nature *496,* 498-503.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. *31,* 227-229.

Keller, P.J. (2013). Imaging morphogenesis: technological advances and biological insights. Science *340,* 1234168.

Lek, M., Dias, J.M., Marklund, U., Uhde, C.W., Kurdija, S., Lei, Q., Sussel, L., Rubenstein, J.L., Matise, M.P., Arnold, H.H., Jessell, T.M., and Ericson, J. (2010). A homeodomain feedback circuit underlies step-function interpretation of a Shh morphogen gradient during ventral neural patterning. Development *137,* 4051-4060.

Megason, S.G., and Fraser, S.E. (2007). Imaging in systems biology. Cell *130,* 784-795.

Paulsson J. (2004). Summing up the noise in gene networks. Nature. *427,* 415-8.

Turing A.M., (1952) The Chemical Basis of Morphogenesis. Phil. Trans. R. Soc. Lond. B *237 641*, 37-72.

Waddington, C.H. (1942). The canalization of development and the inheritance of acquired characteristics. Nature *150,* 563-5.

Wijgerde, M., McMahon, J.A., Rule, M., and McMahon, A.P. (2002). A direct requirement for Hedgehog signaling for normal specification of all ventral progenitor domains in the presumptive mammalian spinal cord. Genes Dev. *16,* 2849-2864.

Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q., *et al.* (2013). TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat. Methods *10,* 329-331.